PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12Q 1/68, C07K 14/47

A1

(11) International Publication Number: WO 97/15686

(43) International Publication Date: 1 May 1997 (01.05.97)

(21) International Application Number: PCT/GB96/02588

(22) International Filing Date: 22 October 1996 (22.10.96)

(30) Priority Data:
60/005,840 23 October 1995 (23.10.95) US

(71) Applicant *(for all designated States except US)*: IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED [GB/GB]; Sardinia House, Sardinia Street, London WC2A 3NL (GB).

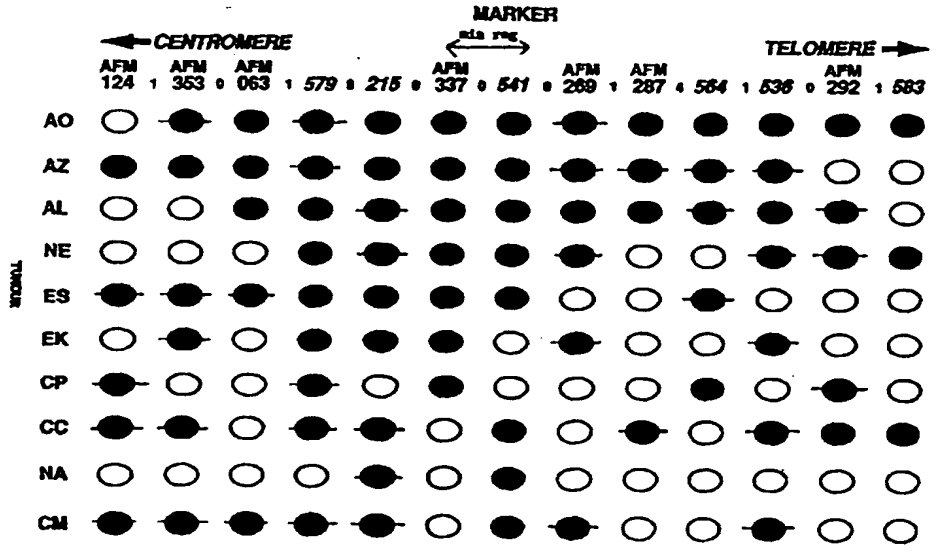
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: SPURR, Nigel, Kay [GB/GB]; 145 Baldock Road, Letchworth, Herts SG6 2EQ (GB). GRAY, Ian, Christopher [GB/GB]; 4 Rycroft Close, Deeping St James, Peterborough PE6 8NX (GB).

(74) Agent: BASSETT, Richard; Eric Potter Clarkson, St. Mary's Court, St. Mary's Gate, Nottingham NG1 1LE (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: DIAGNOSIS OF SUSCEPTIBILITY TO CANCER AND TREATMENT THEREOF



(57) Abstract

A method for determining the susceptibility of a patient to cancer comprising the steps (i) obtaining a sample containing nucleic acid derived from the patient; and (ii) contacting the said nucleic acid with a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215. A nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215 provided that the nucleic acid is not any one of certain YACs, BACs, PACs or ESTs defined herein. Preferably the said nucleic acid is a prostate tumour suppressor gene.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

DIAGNOSIS OF SUSCEPTIBILITY TO CANCER AND TREATMENT THEREOF

The present invention relates to methods of determining whether a patient has cancer or is susceptible to cancer, and it relates to methods of treating cancer, particularly prostate cancer.

Carcinoma of the prostate has become a most significant disease in many countries. Over the last 20 years the mortality rates have doubled and it is now the second commonest cause of male cancer deaths in England and Wales (Mortality Statistics: Cause England and Wales. OPCS DH2 19, 1993, Her Majesty's Stationery Office). The prevalence of prostate cancer has increased by 28% in the last decade and this disease now accounts for 12% of the total cancers of men in England and Wales (Cancer Statistics: Registrations England and Wales. OPCS MBI No 22, 1994, Her Majesty's Stationery Office). This increase and the recent deaths of many public figures from prostatic cancer have served to highlight the need to do something about this cancer. It has been suggested that the wider availability of screening may limit mortality from prostate cancer.

Prostate cancer screening currently consists of a rectal examination and measurement of prostate specific antigen (PSA) levels. These methods lack specificity as digital rectal examination has considerable inter-examiner variability (Smith & Catalona (1995) *Urology* **45**, 70-74) and PSA levels may be elevated in benign prostatic hyperplasia (BPH), prostatic inflammation and other conditions. The comparative failure of PSA as a diagnostic test was shown in 366 men who developed prostate cancer while being included in the Physicians Health Study, a prospective study of over 22,000 men. PSA levels were measured in serum, which

was stored at the start of the study, and elevated levels were found in only 47% of men developing prostate cancer within the subsequent four years (Gann *et al* (1995) *JAMA* 273, 289-294).

Present screening methods are therefore unsatisfactory.

Cytogenetic and allele loss studies have pointed to a number of chromosomal regions of potential involvement in prostate cancer. Cannon-Albright & Eeles (1995) *Nature Genetics* 9, 336-338 (Reference 1) discuss candidate regions for tumour suppressor prostate cancer susceptibility loci from loss-of-heterozygosity (LOH) studies which occur on human chromosome regions 3p, 7q, 8p, 9q, 10p, 10q, 11p, 13q, 16q, 17p, 18q and Y; whereas Brothman *et al* (1990) *Cancer Res.* 50 3795-3803 surveyed cytogenetic information on human prostate adenocarcinoma which indicated loss of chromosomes 1, 2, 5 and Y and gain of 7, 14, 20 and 22, with rearrangements involving chromosome arms 2p, 7q and 10q being most common. Studies by Gao *et al* (1994) *Oncogene* 9, 2999-3003 indicate that a positive mutator phenotype in at least one of chromosomes 3p, 5q, 6p, 7p, 8p, 10q, 11p, 13q, 16q, 17p, 18q and Xq is found in prostate adenocarcinoma; and Massenkeil *et al* (1994) *Anticancer Res.* 14(6B), 2785-2790 indicates that LOH was observed at 8p, 17p, 18q in various prostate tumour samples but no deletions were observed on 10q in fourteen informative prostate tumours. Zenklusen *et al* (1994) *Cancer Res.* 54, 6370-6373 suggests that there is a possible tumour suppressor gene at 7q31.1. In addition, there have been other reports which describe other chromosome loss or abnormalities.

Thus, loss of, or abberations in, most human chromosomes has been implicated in prostate cancer by one research group or another.

A number of tumours exhibit precise loss of the region 10q23-q25 (2, 3), suggesting the presence of a tumour suppressor gene in this area. *Mxi1*, which encodes a negative regulator of the Myc oncoprotein and resides at 10q25, has been proposed as a candidate for the tumour suppressor gene (4); potentially disabling mutations of *Mxi1* in a number of prostate tumours have recently been described. *Mxi1* displays allelic loss and mutation in some cases of prostate cancer and it has been concluded that it may contribute to the pathogenesis or neoplastic evolution of this common malignancy (5).

Objects of the invention are to provide better methods for the diagnosis of cancer and for determining susceptibility to cancer, especially prostate cancer; to provide nucleic acids which are useful in such methods; and to provide a tumour suppressor gene associated with prostate cancer.

Summary of the invention

Using fluorescence based allelotyping with highly informative microsatellite CA repeat markers, we have generated a detailed deletion map spanning 10q23-q25, allowing stricter definition of the region of 10q loss likely to be involved in tumour advancement. In addition, we have assessed the frequency of loss and mutation of *Mxi1* in prostate tumours in order to clarify the role of this gene in prostate tumour progression.

Our data indicate the presence of a prostate tumour suppressor gene (or genes) near the 10q23-q24 boundary, which was deleted in the overwhelming majority (22/23) of tumours showing loss. In contrast, specific loss of *Mxi1*, as opposed to loss of other 10q23-q25 regions or of the entire region, was observed in only 1/23 tumours, and was accompanied by loss of markers at the 10q23-q24 boundary.

Furthermore, we failed to detect any mutations in *Mxi1* in those tumours showing *Mxi1*-associated marker loss by either single-strand conformation polymorphism (SSCP) analysis or direct DNA sequencing, and our data indicate that *Mxi1* is 20 centiMorgans away from the area of chromosome 10 that we have identified. We have found that all tumours which have a loss of 10q have loss of the region specified below.

A first aspect of the invention provides a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215 provided that the nucleic acid is not any one of the yeast artificial chromosomes (YACs) 746-H-8, 821-D-2, 831-E-5, 921-F-8, 738-B-12, 796-D-5, 829-E-1, 678-F-1, 839-B-1, 734-B-4, 7B-F12, 757-D-8, 773-C-2, 787-D-7, 831-E-9, 855-D-2, 855-G-4, 876-G-11, 894-H-5, 922-E-6, 934-D-3, 964-A-8, 968-E-6 or 24G-A10 and is not any one of the expressed sequence tags (ESTs) as described in Tables 3 to 22, and is not any one of the bacterial artificial chromosomes (BACs) or P1-derived artificial chromosomes (PACs) B2F20, P40F10, P72G8, P74N2, P274D21, B76I10, B79A19, B7901, B93F12, B122L22, P201J8, P201P5, P209K3, P316N14, B46B12, B60C5, B145C22, B150K4, B150N3, B181F15, and 188L22.

The position of various markers on human chromosome 10, including D10S541 and D10S215, is as defined in Figure 5. When we refer to these ESTs we mean the sequence that is disclosed in the referenced Tables, and more particularly the specific cDNA clones from which the sequence is derived.

By "selectively hybridising" we mean that the nucleic acid has sufficient nucleotide sequence similarity with the said chromosome 10 DNA that it

can hybridise under moderately or highly stringent conditions. As is well known in the art, the stringency of nucleic acid hybridization depends on factors such as length of nucleic acid over which hybridisation occurs, degree of identity of the hybridizing sequences and on factors such as temperature, ionic strength and CG or AT content of the sequence.

Nucleic acids which can selectively hybridise to the said chromosome 10 DNA include nucleic acids which have >95% sequence identity, preferably those with >98%, more preferably those with >99% sequence identity, over at least a portion of the nucleic acid with the said chromosome 10 DNA. As is well known, human genes usually contain introns such that, for example, a mRNA or cDNA derived from a gene within the said chromosome 10 DNA would not match perfectly along its entire length with the said chromosome 10 DNA but would nevertheless be a nucleic acid capable of selectively hybridising to the said region of chromosome 10.

Typical moderately or highly stringent hybridisation conditions which lead to selective hybridisation are known in the art, for example those described in *Molecular Cloning, a laboratory manual*, 2nd edition, Sambrook *et al* (eds), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.

An example of a typical hybridisation solution when a nucleic acid is immobilised on a nylon membrane and the probe nucleic acid is ≥ 500 bases or base pairs is:

6 x SSC (saline sodium citrate)
0.5% sodium dodecyl sulphate (SDS)
100 μg/ml denatured, fragmented salmon sperm DNA

The hybridisation is performed at 68°C. The nylon membrane, with the nucleic acid immobilised, may be washed at 68% in 1 x SSC or, for high stringency, 0.1 x SSC.

20 x SSC may be prepared in the following way. Dissolve 175.3 g of NaCl and 88.2 g of sodium citrate in 800 ml of H_2O. Adjust the pH to 7.0 with a few drops of a 10 N solution of NaOH. Adjust the volume to 1 litre with H_2O. Dispense into aliquots. Sterilize by autoclaving.

An example of a typical hybridisation solution when a nucleic acid is immobilised on a nylon membrane and the probe is an oligonucleotide of between 15 and 50 bases is:

3.0 M trimethylammonium chloride (TMACl)
0.01 M sodium phosphate (pH 6.8)
1 mm EDTA (pH 7.6)
0.5% SDS
100 μg/ml denatured, fragmented salmon sperm DNA
0.1% nonfat dried milk

The optimal temperature for hybridization is usually chosen to be 5°C below the T_i for the given chain length. T_i is the irreversible melting temperature of the hybrid formed between the probe and its target sequence. Jacobs *et al* (1988) *Nucl. Acids Res.* **16**, 4637 discusses the determination of T_is. The recommended hybridization temperature for 17-mers in 3 M TMACl is 48-50°C; for 19-mers, it is 55-57°C; and for 20-mers, it is 58-66°C.

By "nucleic acid capable of selectively hybridising" we also include nucleic acids which will amplify DNA from the said region of

chromosome 10 by any of the well known amplification systems such as those described in more detail below, in particular the polymerase chain reaction (PCR). Suitable conditions for PCR amplification include amplification in a suitable 1 x amplification buffer:

10 x amplification buffer is 500 mM KCl; 100 mM Tris.Cl (pH 8.3 at room temperature); 15 mM $MgCl_2$; 0.1% gelatin.

Suitably, the annealing part of the amplification is between 37°C and 60°C, preferably 50°C.

The markers D10S541 and DS10S215 define regions on chromosome 10 which are indicated, for example, on the 1993-1994 Genethon human genetic linkage map which is described by Gyapay *et al* (1994) *Nature Genetics* 7, special issue No. 2, 246-339.

The aforementioned YACs are all publicly available from the CEPH mega-YAC library or the ICI YAC library (7B-F12 and 24G-A10), or from the Human Genome Mapping Project Resource Centre, Hinxton Hall, Hinxton, Cambridgeshire, CB10 1RQ, UK. The position of the YACs on the genetic linkage map is made by reference to the CEPH-Genethon Quickmap database (Cohen *et al* (1993) *Nature* 366, 698-701). Sequences of the aforementioned expressed sequence tags (ESTs) are given in Tables 3 to 22 and these are publicly available from GenBank, National Center for Biotechnology Information, National Library of Medicine, Bldg 38A, National Institutes of Health, Rockville Pike, Bethesda, MD 20894, USA. As is described in more detail below, an especially preferred nucleic acid of the invention is a nucleic acid capable of hybridising to the gene corresponding to the cDNA insert of clone IMAGE 264611.

IMAGE clone 264611 is publicly available from Research Genetics, Inc (2130 Memorial Parkway, SW Huntsville, AL 35801, USA) and other IMAGE sources eg American Type Culture Collection, Rockville, MD 20852, USA; Genome Systems Inc, 8629 Pennell Drive, St Louis, Missouri, MO 63114, USA, UK-HGMP Resource Centre, Hinxton, Cambridge CB10 1SB. The clone was obtained as described in the enclosed information for the ESTs N29304 and N20238 (see Tables 9 and 10). The clone is in a modified Pharmacia pT7T3 vector.

NAME: pT7T3D-Pac (ampicillin resistant; 50 μg/ml)
HOST: DH10B
V_TYPE: plasmid
POLYLINKER SEQUENCE: (modified)
tttaatacgactcactatagggaatttggccctcgaggccaagaattcccgactacgtag
tcggggatccgtcttaattaagcggccgcaagcttattccctttagtgagggttaatttt
agcttggcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaa
cttaatcgccttgcagcacatccccctttcgccagctggcgtaatagcgaagag

The sequence of the insert of IMAGE clone 264611 is given in Figure 6.

The following clones contain sequence that is part of the same gene as IMAGE clone 264611 since they overlap to form a largely contiguous sequence. All clones are freely available as physical entities unless otherwise noted. For each clone, some sequence, usually from the 5′ or 3′ ends, is available as ESTs which can be used to produce probes as described below.

The clones and their ESTs are listed on GenBank and the EMBL databases.

EST	cDNA clone	Table No
AA009519	IMAGE 365465 (5')	3
AA009520	IMAGE 365465 (3')	4
AA017563	IMAGE 361374 (3')	5
C01084	-	6
H92038	IMAGE 221326 (5')	7
H92039	IMAGE 221326 (3')	8
N20238	IMAGE 264611 (3')	9
N29304	IMAGE 264611 (5')	10
N35389	IMAGE 272092 (3')	11
N48030	IMAGE 272092 (5')	12
R06763	IMAGE 126556 (3')	13
R06814	IMAGE 126556 (5')	14
R29457	F1-578D (5')	15
T05157	HFBCS42	16
T60214	IMAGE 81420 (5')	17
W23656	IMAGE 306632	18
W27533	-	19
W30684	IMAGE 309597 (5')	20
W81026	IMAGE 347316 (5')	21
W81062	IMAGE 347316 (3')	22

It is preferred if the nucleic acid is capable of selectively hybridising to the region of chromosome 10 bounded by DNA defined by the markers D10S541 and AFM337xf9. Information on the marker AFM337xf9 is freely available from Genethon, 1 rue de L'Internationale, 91000 Evry, France. AFM337xf9 is now known as D10S1765.

It is particularly preferred if the nucleic acid is capable of selectively hybridising to the human-derived DNA of any one of the YACs 746-H-8,

821-D-2, 831-E-5, 921-F-8, 796-D-5, 829-E-1, 839-B-1, 734-B-4 or 24G-A10; and it is still more preferred if the nucleic acid is capable of selectively hybridising to the human-derived DNA of any one of the YACs 746-H-8, 921-F-8, 821-D-2, 831-E-5, 796-D-5, 24G-A-10 or 734-B-4. It will be appreciated that a YAC contains DNA which is required for propagation and maintenance in yeast. The preferred nucleic acids of the invention are those that selectively hybridise to the human-derived DNA present in the YAC and not other DNA in the YAC, such as yeast DNA.

The human-derived cDNA insert of IMAGE clone 264611 hybridises to at least YAC clones 921F8, 746H8, 821D2, 831E5, 796D5 and 24GA10.

The human-derived cDNA insert of IMAGE clone 264611 hybridises to at least BAC (bacterial artificial chromosome) clones B2F20, B46B12, B60C5, B150K4, B150N3, B145C22, B181F15, and B188L22, but not to B76I10, B79A19, B7901, B93F12 and B122L22.

BAC clones are publicly available from Research Genetics, 2130 Memorial Parkway, SW Huntsville, AL 35801, USA and Genome Systems Inc, 8629 Pennell Drive, St Louis, Missouri, MO 63114, USA.

The human-derived cDNA insert of IMAGE clone 264611 hybridises to at least PAC (P1-derived artificial chromosome) clones P40F10 and P274D21, but not to P72G8, P74N2, P201J8, P201P5, P209K3 and P316N14.

The PAC clones are publicly available from the Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.

Although the nucleic acid of the invention may be RNA or DNA, DNA

is preferred. Although the nucleic acid of the invention may be double-stranded or single-stranded, single-stranded nucleic acid is preferred.

The nucleic acid of the invention may be very large, such as 100 kb, if it is double stranded. Indeed genes, such as a tumour suppressor gene, are often this large. However, for diagnostic, probing or amplifying purposes, it is preferred if the nucleic acid has fewer than 10 000, more preferably fewer than 1000, more preferably still from 10 to 100, and in further preference from 15 to 30 base pairs (if the nucleic acid is double-stranded) or bases (if the nucleic acid is single stranded). As is described more fully below, single-stranded DNA primers, suitable for use in a polymerase chain reaction, are particularly preferred.

An especially preferred nucleic acid of the invention is a nucleic acid capable of hybridising to the gene corresponding to the cDNA insert of clone IMAGE 264611 from which EST sequences N29304 and N20238 are derived. The sequence and information for N48030 and N20238 are recorded in the GenBank and EMBL databases (see Tables 9 and 12). Fragments and variants of this gene, and cDNAs derivable from the mRNA encoded by the gene are also preferred nucleic acids of the invention. By "gene corresponding to the cDNA insert clone IMAGE 264611" we mean the gene which encodes mRNA which, when copied in part, produced the cDNA insert in said clone.

Clearly the gene itself and variants and fragments thereof are a preferred nucleic acid of the invention. By "gene" we include not only the introns and exons but also regulatory regions associated with, and physically close to, the introns and exons, particularly those 5′ to the 5′-most exon.

By "fragment" of a gene we include any portion of the gene of at least 15

nucleotides in length (whether single stranded or double stranded) but more preferably the fragment is at least 20 nucleotides in length, most preferably at least 50 nucleotides in length and may be at least 100 nucleotides in length or may be at least 500 nucleotides in length. Preferably the fragment is no more than 50 kb and, more preferably, no more than 100 kb.

By "variant" of a gene we include specifically a cDNA, whether partial or full length, or whether copied from any splice variants of mRNA. We also include specifically a nucleic acid wherein, compared to the natural gene, nucleotide substitutions (including inversions), insertions and deletions are present whether in the gene or a fragment thereof or in a cDNA. Both variants and fragments will be selected according to their intended purposes; for probing, amplifying or diagnostic purposes, shorter fragments but a greater degree of sequence identity (eg at least 80%, 90%, 95% or 99%) will generally be required than for the purposes of expressing a therapeutically useful product, where longer fragments will generally be needed but advantage can be taken of the redundancy in the genetic code, if desired.

It is particularly preferred if the nucleic acid of the invention is an oligonucleotide primer which can be used to amplify a portion of the gene corresponding to the cDNA insert of clone IMAGE 264611.

It is also preferred if the nucleic acid of the invention comprises all or part of the gene and can be used as a probe for hybridisation.

A cDNA sequence of IMAGE 264611 is shown in Figure 1.

The gene and further cDNAs derivable from the gene are readily obtained

using methods well known in the art. For example, further cDNAs can be isolated from a prostate cDNA library using standard methods and the IMAGE 264611 clone as a probe or other probes readily derived from the sequences given in Tables 1 to 19 and the Figures. The sequence is readily determined using standard methods. Similarly, the gene can be isolated from a human genomic DNA library, using the IMAGE 264611 clone as a probe using standard methods or other probes readily derived from the sequences in Tables 1 to 19 and the Figures.

A prostate cDNA library may be obtained using standard molecular biology methods or may be obtained from Clontech Laboratories, Inc, 1020 East Meadow Circle, Palo Alto, California 94303-4230, USA.

Standard methods of screening DNA libraries, isolating and manipulating cloned DNA and sequencing DNA are described in Sambrook *et al* (1989) "Molecular cloning, a laboratory manual", 2nd Edition, Ed Sambrook *et al*, Cold Spring Harbor Press, Cold Spring Harbor, New York.

The predicted amino acid sequence encoded by the IMAGE clone 264611 or the nucleotide sequences shown in Tables 3 to 22 may be used to make peptides which can, in turn, be used to make antibodies. The antibodies can be used to screen a cDNA expression library or can be used to isolate the polypeptide encoded by the gene. Once the polypeptide is isolated its N-terminal sequence can be obtained using methods well known in the art. The amino acid sequence is then used to design an oligonucleotide probe which identifies the 5' coding region of a cDNA.

It will be appreciated that the 5' ends of cDNAs can be isolated by RACE (Rapid Amplification of cDNA Ends; Schaefer (1995) *Anal. Biochem.* 227, 255-273), a technique well known in the art. This approach, and

related approaches, involve reverse transcription from mRNA using a primer based on the presently known 5′ sequence which works back towards the 5′ end of the mRNA transcript followed by PCR using random primers to prime from the "unknown" 5′ end. Messenger RNA-based RACE can also be used for obtaining 5′ ends by isolating mRNA, removing the 5′ cap and then the 5′ end is ligated to an adaptor sequence and PCR follows using one primer against the adaptor and one primer specific to the cDNA of interest.

Methods for isolating genes and parts of genes are described in *Current Protocols in Human Genetics*, 1996, Dracopoli *et al* (ed), John Wiley & Sons, incorporated herein by reference. One useful technique is "vectorette" PCR.

Vectorette PCR can be used for the identification of novel genes, or for the identification of additional sequence when part of the sequence of a gene is already known. The vectorette itself is a double stranded piece of synthetic DNA, with a mismatched central region and one end suitable for ligation to DNA cut by a restriction enzyme (described in *Current Protocols in Human Genetics* 1995 (see pages 5.9.15-5.9.21) and in Valdes *et al* (1994) *Proc. Natl. Acad. Sci. USA* 91, 5377-5381 and Allen *et al PCR Methods and Applications* 4, 71-75). Following ligation of the vectorette to restriction fragments derived from an appropriate DNA source (usually a large genomic DNA fragment such as a YAC clone), PCR amplification is performed using a primer derived from the target DNA in conjunction with a primer derived from the mismatched region of the vectorette. This vectorette primer has the same sequence as the bottom strand of this mismatched region and therefore has no complementary sequence to anneal to in the first cycle of PCR. The first round of amplification is unidirectional, as priming can only occur from

the primer within the target DNA. This produces a complementary strand for the vectorette PCR primer to anneal to in the second PCR cycle. In the second and subsequent cycles of PCR, both primers can prime DNA synthesis with the end result being that the only fragment amplified contains the sequence of interest.

This technique can be used for the identification of intronic sequences within a gene based on a knowledge of the cDNA sequence for that gene. Following restriction digestion of a genomic DNA fragment bearing the gene of interest (such as a YAC clone) and subsequent ligation to the vectorette, a primer designed from the cDNA sequence is used in conjunction with the vectorette primer to PCR amplify a specific fragment of the gene. Exon/intron boundaries can be identified by comparison of the sequence of this fragment to that of the cDNA. This method has been used in combination with primers derived from cDNA clone 264611 to identify intron sequences (see Figures 8-15).

Similarly, a vectorette approach can be used to identify the missing 5′ end of a gene by using a primer derived from the 5′ end of the known cDNA sequence to generate further 5′ sequence data.

Vectorettes can also be used for the identification of completely novel gene sequences in a technique known as 'island rescue'. This approach exploits the fact that CpG-rich 'islands' exist within mammalian genomes and that such islands are associated with the 5′ ends of genes. Certain restriction enzymes cut within CpG islands, for example, the enzyme *Not*I. Following *Not*I digestion of a genomic DNA fragment, a vectorette with a *Not*I-compatible sticky end is ligated to the resulting sub-fragments. PCR amplification is then performed using the vectorette primer in conjunction with a primer derived from an *Alu* repeat element. Such

elements occur at frequent intervals in the human genome, therefore it is likely that one or more will lie adjacent to the CpG island of interest and facilitate the generation of a PCR product. As a control, a second PCR reaction is executed, excluding the vectorette primer. Any fragments generated in the *Alu*/vectorette primed reaction but absent from the *Alu* only control should represent part of the CpG island and can be gel-purified and analysed for coding sequences using standard methods.

The polypeptide encoded by the gene corresponding to the cDNA clone IMAGE 264611 or the nucleotide sequences shown in Tables 3 to 22 has some sequence similarity to the polypeptide tensin, a protein involved in cytoskeletal/extracellular matrix interactions; similarity is also observed, at least at the nucleotide sequence level, with auxilin, a protein involved in protein transport to the cell membrane *via* clathrin coated vesicles. Sequence similarity between tensin and auxilin has also been noted previously.

A preferred nucleic acid of the invention is one comprising a tumour suppressor gene or fragment or variant thereof. The tumour suppressor gene is one which is involved in the origin or development of a cancer such as prostate cancer, melanoma, glioma or non-Hodgkin's lymphoma. Suitably, the tumour suppressor gene is involved in the origin or development of prostate cancer, particularly prostate adenocarcinoma.

A nucleic acid of the invention comprising a tumour suppressor gene or fragment or derivative thereof is readily identified; for example, the gene may be identified by screening a panel of RNAs from prostate and other tumour cell lines in order to identify a reduced level of transcript. The transcript may be large, as it will probably have a complex function and several sites for disabling mutation 'hits' (as is the case with the tumour

suppressor genes BRCA1, RB). Cross-species conservation indicates that the gene has a basic cell 'housekeeping' function, the loss of which may lead to loss of growth control and tumour formation.

By "tumour suppressor gene" we include any gene for which loss or some reduction in any of its function or activities can contribute to neoplasia.

Analysis of the entire coding region of the tumour suppressor gene in tumours indicates that the gene is a tumour suppressor gene when the gene has been altered compared to the gene in non-tumour tissue or to the gene in an individual who does not have, and who is not prone to, prostate cancer, and that it is involved in the cancer, such as prostate cancer. Suitable methods for mutation analysis include single-stranded conformation polymorphism (SSCP) analysis (or variations of this technique) and direct DNA sequencing. These are well known to the person skilled in the art, and SSCP, for example, is described in *Current Protocols in Human Genetics*, 1995, pp 7.4.1-7.4.6.

Any tumour suppressor gene of the invention almost certainly contains introns (as does the gene corresponding to IMAGE clone 264611) and almost certainly is >0.5 kb, more likely >1.0 kb and most likely between 1.0 kb and 500 kb. The cDNA insert in IMAGE clone 264611 is about 1.7 kbp. Any tumour suppressor gene of the invention almost certainly is polymorphic in its DNA sequence. Thus, fragments (such as restriction fragments or fragments derived by enzymatic amplification) and variants (such as natural variants, eg allelic variants) or variants created by *in vitro* manipulation are part of the invention. Suitable such fragments include fragments which are useful as a hybridisation probe or fragments which are useful as an amplification primer. Suitable such variants include variants in which the coding sense of the gene is unaltered or

variants in which the coding sequence is modified so as to alter the properties of the encoded polypeptide.

Although any tumour suppressor gene of the invention almost certainly ultimately encodes a polypeptide, it may encode an RNA species which RNA species does not encode a polypeptide.

It is further preferred if the nucleic acid comprises a nucleic acid product of a tumour suppressor gene or derivative or fragment or variant thereof. Such nucleic acids include mRNA transcribed from the tumour suppressor gene.

It is particularly preferred if the nucleic acid is a cDNA (copy DNA) derived from a mRNA transcribed from the tumour suppressor gene. Libraries of cDNA derived from selected tissues, such as prostate or prostate tumour tissue, are known in the art and can be prepared from suitable mRNA using methods known in the art for example as described in *Molecular cloning, a laboratory manual* (*supra*).

The nucleotide sequences described in Tables 3 to 22 are partial sequences of partial cDNAs the said cDNAs being derived from mRNAs which are related to, selectively hybridise to, and are almost certainly transcribed from a gene or genes found in the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215. The nucleotide sequences shown in Figures 8 to 15 include sequences from introns in the gene corresponding to IMAGE clone 264611. More particularly, we have found that polynucleotides comprising the sequences of any of Tables 3 to 22 and Figures 6 and 8 to 15 hybridise to at least one of the aforementioned YAC, BAC and PAC clones. Thus, the nucleotide sequences of Tables 3 to 22 and Figure 6

represent the mRNA products of at least one gene which is found within the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215; more particularly in the sub-region defined by the YAC clones. A particularly preferred embodiment comprises a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215 and capable of selectively hybridising to the human-derived sequence as described in any one of Tables 3 to 22 and Figures 6 and 8 to 15 provided that the nucleic acid is not any one of the yeast artificial chromosomes (YACs) 746-H-8, 821-D-2, 831-E-5, 921-F-8, 738-B-12, 796-D-5, 829-E-1, 678-F-1, 839-B-1, 734-B-4, 7B-F12, 757-D-8, 773-C-2, 787-D-7, 829-E-1, 831-E-9, 855-D-2, 855-G-4, 876-G-11, 894-H-5, 921-F-8, 922-E-6, 934-D-3, 964-A-8, 968-E-6 or 24G-A10 and is not any one of the polynucleotides as described in Tables 3 to 22 and is not any one of the BACs or PACs B2F20, P40F10, P72G8, P74N2, P274D21, B76I10, B79A19, B7901, B93F12, B122L22, P201J8, P201P5, P209K3, P316N14, 46B12, B60C5, B145C22, B150K4, B150N3, B181F15, and B188L22.

It will readily be appreciated that a person skilled in the art can identify a gene or genes which correspond to IMAGE clone 264611 by making use of the sequence information presented in Tables 3 to 22 and Figures 6 and 8 to 15.

In particular, it is preferred if the nucleic acid comprises the gene or genes from which the sequence of any one of Tables 3 to 22 and Figures 6 and 8 to 15 is derived or a fragment or variant thereof. It is also preferred if the nucleic acid comprises a full length cDNA or a cDNA which is at least 50% of the length of a mRNA transcript; more preferably greater than 75% of the length; more preferably greater than 95% of the length.

It may be desirable to subclone the nucleic acid, particularly if all or part of the protein coding sequence is to be expressed.

Generally, the DNA is inserted into an expression vector, such as a plasmid, in proper orientation and correct reading frame for expression. If necessary, the DNA may be linked to the appropriate transcriptional and translational regulatory control nucleotide sequences recognised by the desired host, although such controls are generally available in the expression vector. The vector is then introduced into the host through standard techniques. Generally, not all of the hosts will be transformed by the vector. Therefore, it will be necessary to select for transformed host cells. One selection technique involves incorporating into the expression vector a DNA sequence, with any necessary control elements, that codes for a selectable trait in the transformed cell, such as antibiotic resistance. Alternatively, the gene for such selectable trait can be on another vector, which is used to co-transform the desired host cell.

Host cells that have been transformed by the recombinant DNA of the invention are then cultured for a sufficient time and under appropriate conditions known to those skilled in the art in view of the teachings disclosed herein to permit the expression of the polypeptide, which can then be recovered.

Many expression systems are known, including bacteria (for example *E. coli* and *Bacillus subtilis*), yeasts (for example *Saccharomyces cerevisiae*), filamentous fungi (for example *Aspergillus*), plant cells, animal cells and insect cells.

The vectors include a prokaryotic replicon, such as the ColE1 *ori*, for propagation in a prokaryote, even if the vector is to be used for expression

in other, non-prokaryotic, cell types. The vectors can also include an appropriate promoter such as a prokaryotic promoter capable of directing the expression (transcription and translation) of the genes in a bacterial host cell, such as *E. coli*, transformed therewith.

A promoter is an expression control element formed by a DNA sequence that permits binding of RNA polymerase and transcription to occur. Promoter sequences compatible with exemplary bacterial hosts are typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention.

Typical prokaryotic vector plasmids are pUC18, pUC19, pBR322 and pBR329 available from Biorad Laboratories, (Richmond, CA, USA) and p*Trc*99A and pKK223-3 available from Pharmacia, Piscataway, NJ, USA.

A typical mammalian cell vector plasmid is pSVL available from Pharmacia, Piscataway, NJ, USA. This vector uses the SV40 late promoter to drive expression of cloned genes, the highest level of expression being found in T antigen-producing cells, such as COS-1 cells.

An example of an inducible mammalian expression vector is pMSG, also available from Pharmacia. This vector uses the glucocorticoid-inducible promoter of the mouse mammary tumour virus long terminal repeat to drive expression of the cloned gene.

Useful yeast plasmid vectors are pRS403-406 and pRS413-416 and are generally available from Stratagene Cloning Systems, La Jolla, CA 92037, USA. Plasmids pRS403, pRS404, pRS405 and pRS406 are Yeast Integrating plasmids (YIps) and incorporate the yeast selectable markers *HIS3*, *TRP1*, *LEU2* and *URA3*. Plasmids pRS413-416 are Yeast

Centromere plasmids (YCps)

A variety of methods have been developed to operably link DNA to vectors via complementary cohesive termini. For instance, complementary homopolymer tracts can be added to the DNA segment to be inserted to the vector DNA. The vector and DNA segment are then joined by hydrogen bonding between the complementary homopolymeric tails to form recombinant DNA molecules.

Synthetic linkers containing one or more restriction sites provide an alternative method of joining the DNA segment to vectors. The DNA segment, generated by endonuclease restriction digestion as described earlier, is treated with bacteriophage T4 DNA polymerase or *E. coli* DNA polymerase I, enzymes that remove protruding, 3′-single-stranded termini with their 3′-5′-exonucleolytic activities, and fill in recessed 3′-ends with their polymerizing activities.

The combination of these activities therefore generates blunt-ended DNA segments. The blunt-ended segments are then incubated with a large molar excess of linker molecules in the presence of an enzyme that is able to catalyze the ligation of blunt-ended DNA molecules, such as bacteriophage T4 DNA ligase. Thus, the products of the reaction are DNA segments carrying polymeric linker sequences at their ends. These DNA segments are then cleaved with the appropriate restriction enzyme and ligated to an expression vector that has been cleaved with an enzyme that produces termini compatible with those of the DNA segment.

Synthetic linkers containing a variety of restriction endonuclease sites are commercially available from a number of sources including International Biotechnologies Inc, New Haven, CN, USA.

Particularly preferred nucleic acids of the first aspect of the invention are those selected from the group consisting of primers suitable for amplifying nucleic acid. Suitably, the nucleic acids are selected from the group consisting of primers which hybridise to the nucleotide sequences as described in any one of Tables 3 to 22 and Figures 6 and 8 to 15, or their complement.

It is particularly preferred if the amplification primers hybridise to the introns of a gene. They are particularly useful if processed pseudogenes are present. Thus, it is preferred if the nucleic acids are selected from the group consisting of primers which hybridise to the sequences given in Figures 6 and 8 to 15, or their complement.

Primers which are suitable for use in a polymerase chain reaction (PCR; Saiki *et al* (1988) *Science* **239**, 487-491) are preferred. Suitable PCR primers may have the following properties:

It is well known that the sequence at the 5′ end of the oligonucleotide need not match the target sequence to be amplified.

It is usual that the PCR primers do not contain any complementary structures with each other longer than 2 bases, especially at their 3′ ends, as this feature may promote the formation of an artifactual product called "primer dimer". When the 3′ ends of the two primers hybridize, they form a "primed template" complex, and primer extension results in a short duplex product called "primer dimer".

Internal secondary structure should be avoided in primers. For symmetric PCR, a 40-60% G+C content is often recommended for both primers, with no long stretches of any one base. The classical melting temperature

calculations used in conjunction with DNA probe hybridization studies often predict that a given primer should anneal at a specific temperature or that the 72°C extension temperature will dissociate the primer/template hybrid prematurely. In practice, the hybrids are more effective in the PCR process than generally predicted by simple T_m calculations.

Optimum annealing temperatures may be determined empirically and may be higher than predicted. *Taq* DNA polymerase does have activity in the 37-55°C region, so primer extension will occur during the annealing step and the hybrid will be stabilized. The concentrations of the primers are equal in conventional (symmetric) PCR and, typically, within 0.1- to 1-μM range.

Any of the nucleic acid amplification protocols can be used in the method of the invention including the polymerase chain reaction, QB replicase and ligase chain reaction. Also, NASBA (nucleic acid sequence based amplification), also called 3SR, can be used as described in Compton (1991) *Nature* **350**, 91-92 and *AIDS* (1993), Vol 7 (Suppl 2), S108 or SDA (strand displacement amplification) can be used as described in Walker *et al* (1992) *Nucl. Acids Res.* **20**, 1691-1696. The polymerase chain reaction is particularly preferred because of its simplicity.

When a pair of suitable nucleic acids of the invention are used in a PCR it is convenient to detect the product by gel electrophoresis and ethidium bromide staining. As an alternative to detecting the product of DNA amplification using agarose gel electrophoresis and ethidium bromide staining of the DNA, it is convenient to use a labelled oligonucleotide capable of hybridising to the amplified DNA as a probe. When the amplification is by a PCR the oligonucleotide probe hybridises to the interprimer sequence as defined by the two primers. The oligonucleotide

probe is preferably between 10 and 50 nucleotides long, more preferably between 15 and 30 nucleotides long. The probe may be labelled with a radionuclide such as ^{32}P, ^{33}P and ^{35}S using standard techniques, or may be labelled with a fluorescent dye. When the oligonucleotide probe is fluorescently labelled, the amplified DNA product may be detected in solution (see for example Balaguer *et al* (1991) "Quantification of DNA sequences obtained by polymerase chain reaction using a bioluminescence adsorbent" *Anal. Biochem.* **195**, 105-110 and Dilesare *et al* (1993) "A high-sensitivity electrochemiluminescence-based detection system for automated PCR product quantitation" *BioTechniques* **15**, 152-157.

PCR products can also be detected using a probe which may have a fluorophore-quencher pair or may be attached to a solid support or may have a biotin tag or they may be detected using a combination of a capture probe and a detector probe.

Fluorophore-quencher pairs are particularly suited to quantitative measurements of PCR reactions (eg RT-PCR). Fluorescence polarisation using a suitable probe may also be used to detect PCR products.

Further particularly preferred nucleic acids are those which will act as PCR primers which primers can be selected by reference to the sequence shown in Figures 6 and 8 to 15. These primers are useful in amplifying DNA derived from the gene corresponding to the cDNA clone IMAGE 264611. These primers include, but are not limited to, the sequences which are given on Figures 8 to 15 in bold (see Figure legends). The downstream (3') primers are the reverse complement of the sequences indicated in bold.

Oligonucleotide primers can be synthesised using methods well known in

the art, for example using solid-phase phosphoramidite chemistry.

A second aspect of the invention provides a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215, further comprising a detectable label.

By "detectable label" we include any convenient radioactive label such as ^{32}P, ^{33}P or ^{35}S which can readily be incorporated into a nucleic acid molecule using well known methods; we also include any convenient fluorescent or chemiluminescent label which can readily be incorporated into a nucleic acid. In addition the term "detectable label" also includes a moiety which can be detected by virtue of binding to another moiety (such as biotin which can be detected by binding to streptavidin); and a moiety, such as an enzyme, which can be detected by virtue of its ability to convert a colourless compound into a coloured compound, or vice versa (for example, alkaline phosphatase can convert colourless o-nitrophenylphosphate into coloured o-nitrophenol). Conveniently, the nucleic acid probe may occupy a certain position in a fixed assay and whether the nucleic acid hybridises to the said region of human chromosome 10 can be determined by reference to the position of hybridisation in the fixed assay. The detectable label may also be a fluorophore-quencher pair as described in Tyagi & Kramer (1996) *Nature Biotechnology* **14**, 303-308.

It is preferred if the nucleic acid comprises the human-derived sequence in any one of the expressed sequence tags (ESTs) as described in Tables 3 to 22 or the cDNA described in Figure 6 or the intron sequences shown in Figures 8 to 15 further comprising a detectable label; or if the nucleic acid comprises the human-derived sequence in any one of the yeast

artificial chromosomes (YACs) 921-F-8, 746-H-8, 821-D-2, 831-E-5, 796-D-5, 24G-A-10 or 734-B-4 or BAC clones B2F20, B46B12, B60C5, B150K4, B150N3, B145C22, B181F15, B188L22, or PAC clones P40F10, and P274D21.

Particularly preferred nucleic acids are those of the first aspect of the invention further comprising a detectable label.

A third aspect of the invention provides a method for determining the susceptibility of a patient to cancer comprising the steps (i) obtaining a sample containing nucleic acid derived from the patient; and (ii) contacting the said nucleic acid with a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215.

The method is suitable for determining the susceptibility of a patient to any cancer but it is preferred if the cancer for which susceptibility is determined is prostate cancer, melanoma, glioma or non-Hodgkin's lymphoma. The method is most suited for determining the susceptibility of a patient to prostate cancer. Accordingly, at least for the determination of susceptibility to prostate cancer, the patient is male.

The presence or absence of a portion of human chromosome 10 may be determined by the methods of the third, fourth and fifth aspects of the invention, and in a preferred embodiment of the third, fourth and fifth aspects of the invention the nucleic acid capable of selectively hybridising to the said region of human chromosome 10 is a nucleic acid suitable for amplification of a portion of the said region of chromosome 10.

A fourth aspect of the invention provides a method of diagnosing cancer

in a patient comprising the steps of (i) obtaining a sample containing nucleic acid derived from the patient; and (ii) contacting the said nucleic acid with a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215.

The method is particularly suited for distinguishing between neoplasia and hyperplasia of the prostate. Because all tumours which have a loss of 10q have also been found to lack the region specified herein, a differential diagnostic test can be performed, using the markers of the invention and other markers (including markers on other chromosomes).

A fifth aspect of the invention provides a method of predicting the relative prospects of a particular outcome of a cancer in a patient comprising the steps of (i) obtaining a sample containing nucleic acid derived from the patient; and (ii) contacting the said nucleic acid with a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215.

Although any sample containing nucleic acid derived from the patient is useful in the methods of the third, fourth and fifth aspects of the invention, it is preferred if the sample is selected from the group consisting of prostate tissue, blood, urine or semen. Prostate tissue can be obtained from a patient using standard surgical techniques. Cells derived from the prostate are found in small numbers in the urine and in the blood. Although it is preferred that the sample containing nucleic acid from the patient is, or is derived directly from, a cell of the patient, such as a prostate cell, a sample indirectly derived from a patient, such as a cell grown in culture, is also included within the invention. Equally, although

the nucleic acid derived from the patient may have been physically within the patient, it may alternatively have been copied from nucleic acid which was physically within the patient. The tumour tissue may be taken from the primary tumour or from metastases, and particularly may be taken from the margins of the tumour.

Conveniently, the nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215 further comprises a detectable label. The detectable label includes the labels described above in relation to the second aspect of the invention.

It will be appreciated that the aforementioned methods may be used for presymptomatic screening of a patient who is in a risk group for cancer. For example, men older than about 60 years are at greater risk of prostate cancer than men below the age of 35. Similarly, the methods may be used for the pathological classification of tumours such as prostate tumours.

Conveniently, in the methods of the third, fourth and fifth aspects of the invention the nucleic acid which is capable of the said selective hybridisation (whether labelled with a detectable label or not) is contacted with a nucleic acid derived from the patient under hybridising conditions. Suitable hybridising conditions include those described in relation to the first aspect of the invention.

It is preferred that if blood, semen or urine is the source of the said sample containing nucleic acid derived from the patient that the sample is enriched for prostate-derived tissue or cells. Enrichment for prostate cells may be achieved using, for example, cell sorting methods such as

fluorescent activated cell sorting (FACS) using a prostate-selective antibody such as one directed to prostate-specific antigen (PSA). The source of the said sample also includes biopsy material and tumour samples, also including fixed paraffin mounted specimens as well as fresh or frozen tissue.

The methods of the third, fourth or fifth aspect of the invention may involve sequencing of DNA at one or more of the relevant positions within the relevant region, including direct sequencing; direct sequencing of PCR-amplified exons; differential hybridisation of an oligonucleotide probe designed to hybridise at the relevant positions within the relevant region (conveniently this uses immobilised oligonucleotide probes in, so-called, "chip" systems which are well known in the art); denaturing gel electrophoresis following digestion with an appropriate restriction enzyme, preferably following amplification of the relevant DNA regions; S1 nuclease sequence analysis; non-denaturing gel electrophoresis, preferably following amplification of the relevant DNA regions; conventional RFLP (restriction fragment length polymorphism) assays; heteroduplex analysis; selective DNA amplification using oligonucleotides; fluorescent *in-situ* hybridisation of interphase chromosomes; ARMS-PCR (Amplification Refractory Mutation System-PCR) for specific mutations; cleavage at mismatch sites in hybridised nucleic acids (the cleavage being chemical or enzymic); SSCP single strand conformational polymorphism or DGGE (discontinuous or denaturing gradient gel electrophoresis); analysis to detect mismatch in annealed normal/mutant PCR-amplified DNA; and protein truncation assay (translation and transcription of exons - if a mutation introduces a stop codon a truncated protein product will result). Other methods may be employed such as detecting changes in the secondary structure of single-stranded DNA resulting from changes in the primary sequence, for example, using the cleavage I enzyme. This system

is commercially available from GibcoBRL, Life Technologies, 3 Fountain Drive, Inchinnan Business Park, Paisley PA4 9RF, Scotland.

Detailed methods of mutation detection are described in "Laboratory Protocols for Mutation Detection" 1996, ed. Landegren, Oxford University Press on behalf of HUGO (Human Genome Organisation).

It is preferred if RFLP is used for the detection of fairly large ($\geq$ 500bp) deletions or insertions. Southern blots may be used for this method of the invention.

PCR amplification of smaller regions (maximum 300bp) to detect small changes greater than 3-4 bp insertions or deletions may be preferred. Amplified sequence may be analysed on a sequencing gel, and small changes (minimum size 3-4 bp) can be visualised. Suitable primers are designed as herein described.

In addition, using either Southern blot analysis or PCR restriction enzyme variant sites may be detected.

For example, for genomic DNA: restriction enzyme digestion, gel electrophoresis, Southern blotting, and hybridisation specific probe (any of the YACs, BACs, in the region as described herein, or a suitable fragment derived therefrom).

For example for PCR: amplify DNA, restriction enzyme digestion, gel detection by ethidium bromide, silver staining or incorporation of radionucleotide or fluorescent primer in the PCR.

Other suitable methods include the development of allele specific

oligonucleotides (ASOs) for specific mutational events. Similar methods are used on RNA and cDNA for prostate specific tissue.

The method also includes checking for loss-of-heterozygosity (LOH; shows one copy lost) and then look for loss of function of RNA by failing to detect a mRNA on Northern blots or by PCR or in RNA/cDNA (shows other copy non-active). LOH on a tumour cell, from whatever source, compared to blood is useful as a diagnostic tool, eg show that the tumour has progressed and requires more stringent treatment.

Preferably, in the third, fourth and fifth aspects of the invention, the nucleic acid is capable of selectively hybridising to the region of human chromosome 10 which region is bounded by the markers D10S541 and D10S215; more preferably the said nucleic acid comprises or is capable of selectively hybridising to the human-derived DNA of any one of YACs 746-H-8, 821-D-2, 831-E-5, 921-F-8, 796-D-5, 829-E-1, 839-B-1, 734-B-4 or 24G-A10; more preferably still the nucleic acid comprises or is capable of selectively hybridising to the human-derived DNA of any one of the YACs 821-D-2, 831-E-5, 796-D-5, 24G-A-10 or 734-B-4.

It is also preferred if the nucleic acid comprises or is capable of selectively hybridising to the human-derived DNA of any of the BACs or PACs B2F20, P40F10, P72G8, P74N2, P274D21, B76I10, B79A19, B7901, B93F12, B122L22, P201J8, P201P5, P209K3, P316N14, B46B12, B60C5, B145C22, B150K4, B150N3, B181F15, and B188L22.

It is also preferred if the nucleic acid is a primer for the microsatellite markers D10S541, D10S215 and AFM337xf9 (D10S1765), namely:
5'-AAGCAAGTGAAGTCTTAGAACCACC-3'
5'-CCACAAGTAACAGAAAGCCTGTCTC-3'

5'-TGGCATCATTCTGGGGA-3'
5'-GCTTTACGTTTCTTCACATGGT-3'
5'-ACACTTACATAGTGCTTTCTGCG-3', and
5'-CAGCCTCCCAAAGTTGC-3'.

It is particularly preferred if the nucleic acid is capable of selectively hybridising to the gene corresponding to the cDNA insert of the clone IMAGE 264611.

Thus, the present invention provides a use of a nucleic acid which is capable of selectively hybridising to the said region of human chromosome 10 in diagnosing cancer or diagnosing susceptibility to cancer.

Also, the present invention provides a method of determining the presence or absence, or a mutation in, the said region of human chromosome 10.

Preferably, the said nucleic acid capable of selectively hybridising is DNA, and also preferably the said nucleic acid is single-stranded.

It is particularly preferred if the said nucleic acid capable of selectively hybridising has fewer than 10 000 base pairs when the nucleic acid is double-stranded or bases when the nucleic acid is single-stranded; more preferably if the said nucleic acid has fewer than 1000 base pairs when the nucleic acid is double-stranded or bases when the nucleic acid is single-stranded; more preferably still if the said nucleic acid has from 10 to 100 base pairs when the nucleic acid is double-stranded or bases when the nucleic acid is single-stranded; and even more preferably if the said nucleic acid has from 15 to 30 base pairs when the nucleic acid is double-stranded or bases when the nucleic acid is single-stranded.

It is preferred if the said nucleic acid capable of selectively hybridising comprises a tumour suppressor gene or fragment or variant thereof, or a nucleic acid which selectively hybridises thereto.

It is preferred if the said nucleic acid capable of selectively hybridising is suitable as a primer for nucleic acid amplification. Suitable primers include those described in relation to the first and second aspects of the invention.

In a preferred embodiment, reverse transcriptase PCR is used to detect micrometastases in blood samples from the patient. A blood sample is taken and RNA prepared from the nucleated cells in the sample. This is used in PCR amplification with oligonucleotide primers detecting the presence or absence, or mutations in prostate tumour suppressor mRNA. This is a relatively sensitive method that can detect one cell in a mix of more than a million normal cells and it is possible to detect prostate tumour suppressor mRNA products present in circulating metastatic cells mixed with normal blood cells that do not express these genes. The gene products of those genes present in the region of chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215, are useful markers detecting circulating prostate cells.

It will be appreciated that it is also possible to detect micrometastases by looking for mutations in the DNA of cells in the blood sample directly, or by using the protein truncation test or by using microsatellite markers; in this case the suspected tumour cells should be purified from the blood.

It is also preferred if the said nucleic acid capable of selectively hybridising is, or is capable of hybridising to, the human derived sequence as described in Tables 3 to 22 or Figures 6 and 8 to 15; conveniently the

said nucleic acid is selected from the group consisting of primers which hybridise to DNA from the sequences as described in Tables 3 to 22 or Figures 6 and 8 to 15.

The methods of the invention include the detection of mutations in the region of chromosome 10 bounded by DNA defined by the markers D10S541 and D10S215; especially in the tumour suppressor gene.

The methods of the invention may make use of a difference in restriction enzyme cleavage sites caused by mutation. A non-denaturing gel may be used to detect differing lengths of fragments resulting from digestion with an appropriate restriction enzyme. The DNA is usually amplified before digestion, for example using the polymerase chain reaction (PCR) method and modifications thereof. Otherwise 10-100 times more DNA would need to be obtained in the first place, and even then the assay would work only if the restriction enzyme cuts DNA infrequently.

Amplification of DNA may be achieved by the established PCR method as disclosed by Saiki *et al* (1988) *Science* **239**, 487-491 or by developments thereof or alternatives such as the ligase chain reaction, QB replicase and nucleic acid sequence-based amplification or other known amplification methods, some of which are described herein.

An "appropriate restriction enzyme" is one which will recognise and cut the wild-type sequence and not the mutated sequence or *vice versa*. The sequence which is recognised and cut by the restriction enzyme (or not, as the case may be) can be present as a consequence of the mutation or it can be introduced into the normal or mutant allele using mismatched oligonucleotides in the PCR reaction. It is convenient if the enzyme cuts DNA only infrequently, in other words if it recognises a sequence which

occurs only rarely.

In another method, a pair of PCR primers are used which match (ie hybridise to) either the wild-type genotype or the mutant genotype but not both. Whether amplified DNA is produced will then indicate the wild-type or mutant genotype (and hence phenotype). However, this method relies partly on a negative result (ie the absence of amplified DNA) which could be due to a technical failure. It is therefore less reliable and/or requires additional control experiments.

A preferable method employs similar PCR primers but, as well as hybridising to only one of the wild-type or mutant sequences, they introduce a restriction site which is not otherwise there in either the wild-type or mutant sequences.

The nucleic acids provided by the present invention are useful for a number of purposes. They can be used in Southern hybridization to genomic DNA and in the RNase protection method for detecting point mutations already discussed above. The probes can be used to detect PCR amplification products. They may also be used to detect mismatches with the tumour suppressor gene or mRNA using other techniques. Mismatches can be detected using either enzymes (eg S1 nuclease or resolvase), chemicals (eg hydroxylamine or osmium tetroxide and piperidine), or changes in electrophoretic mobility of mismatched hybrids as compared to totally matched hybrids. These techniques are known in the art. Generally, the probes are complementary to the tumour suppressor gene coding sequences, although probes to certain introns are also contemplated. An entire battery of nucleic acid probes may be used to compose a kit for detecting loss of or mutation in wild-type tumour suppressor genes. The kit allows for hybridization to the entire tumour

suppressor gene. The probes may overlap with each other or be contiguous.

If a riboprobe is used to detect mismatches with mRNA, it is complementary to the mRNA of the human wild-type tumour suppressor gene. The riboprobe thus is an anti-sense probe in that it does not code for the protein encoded by the tumour suppressor gene because it is of the opposite polarity to the sense strand. The riboprobe generally will be labelled, for example, radioactively labelled which can be accomplished by any means known in the art. If the riboprobe is used to detect mismatches with DNA it can be of either polarity, sense or anti-sense. Similarly, DNA probes also may be used to detect mismatches.

Nucleic acid probes may also be complementary to mutant alleles of the tumour suppressor gene. These are useful to detect similar mutations in other patients on the basis of hybridization rather than mismatches. As mentioned above, the tumour suppressor gene probes can also be used in Southern hybridizations to genomic DNA to detect gross chromosomal changes such as deletions and insertions. The probes can also be used to select cDNA clones of tumour suppressor genes from tumour and normal tissues. In addition, the probes can be used to detect tumour suppressor gene mRNA in tissue to determine if expression is altered, for example diminished, as a result of loss of wild-type tumour suppressor genes.

According to the diagnostic and prognostic method of the present invention, loss of the wild-type gene is detected. The loss may be due to either insertional, deletional or point mutational events. If only a single allele is mutated, an early neoplastic state may be indicated. However, if both alleles are mutated then a malignant state is indicated or an increased probability of malignancy is indicated. The finding of such mutations thus

provides both diagnostic and prognostic information. A tumour suppressor gene allele which is not deleted (eg that on the sister chromosome to a chromosome carrying a gene deletion) can be screened for other mutations, such as insertions, small deletions, and point mutations. It is believed that most mutations found in tumour tissues will be those leading to greatly altered expression of the tumour suppressor gene product. However, mutations leading to non-functional gene products would also lead to a malignant state or an increased probability of malignancy. Mutational events (such as point mutations, deletions, insertions and the like) may occur in regulatory regions, such as in the promoter of the gene, leading to loss or diminution of expression of the mRNA. Point mutations may also abolish proper RNA processing, leading to loss of expression of the tumour suppressor gene product.

The invention also includes the following methods: *in vitro* transcription and translation of tumour suppressor gene to identify truncated gene products, or altered properties such as substrate binding; immunohistochemistry of tissue sections to identify cells in which expression of the protein is reduced/lost, or its distribution is altered within cells or on their surface; and the use of RT-PCR using random primers, prior to detection of mutations in the region as described above.

A sixth aspect of the invention provides a system (or it could also be termed a kit of parts) for detecting the presence or absence of, or mutation in, the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215, the system comprising a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215 and a nucleoside triphosphate or deoxynucleoside triphosphate or derivative thereof. Preferred nucleic acids capable of

selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by markers D10S541 and D10S215 are the same as those preferred in the third, fourth and fifth aspects of the invention.

By "mutation" we include insertions, substitutions and deletions.

By "nucleoside triphosphate or deoxynucleoside triphosphate or derivative thereof" we include any naturally occurring nucleoside triphosphate or deoxynucleoside triophosphate such as ATP, GTP, CTP, and UTP, dATP dGTP, dCTP, TTP as well as non-naturally derivatives such as those that include a phosphorothioate linkage (for example αS derivatives).

Conveniently the nucleoside triphosphate or deoxynucleoside triphosphosphate is radioactively labelled or derivative thereof, for example with ^{32}P, ^{33}P or ^{35}S, or is fluorescently labelled or labelled with a chemiluminescence compound or with digoxygenin.

Conveniently deoxynucleotides are at a concentration suitable for dilution to use in a PCR.

Thus, the invention includes a kit of parts which includes a nucleic acid capable of selectively hybridising to the said region of human chromosome 10 and means for detecting the presence or absence of, or a mutation in, the said region.

A seventh aspect of the invention provides a system for detecting the presence or absence of, or mutation in, the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215, the system comprising a nucleic acid capable of selectively

hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215 and a nucleic acid modifying enzyme. Preferred nucleic acids capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by markers D10S541 and D10S215 are the same as those preferred in the third, fourth and fifth aspects of the invention.

By "mutation" we include insertions, substitutions (including transversions) and deletions.

By "nucleic acid modifying enzyme" we include any enzyme capable of modifying an RNA or DNA molecule.

Preferred enzymes are selected from the group consisting of DNA polymerases, DNA ligases, polynucleotide kinases or restriction endonucleases. A particularly preferred enzyme is a thermostable DNA polymerase such as *Taq* DNA polymerase. Nucleases such as Cleavase I which recognise secondary structure, for example mismatches, may also be useful.

An eighth aspect of the invention provides a polypeptide capable of being encoded by the tumour suppressor gene of the invention or a fragment or variant thereof. The polypeptide preferably has tumour suppressor activity, especially in the prostate, or cross-reacts with an antibody which is specific for the native polypeptide.

A ninth aspect of the invention comprises a molecule capable of specifically binding with a polypeptide of the eighth aspect of the invention. Suitably, the molecule is an antibody-like molecule comprising

complementarity-determining regions specific for the said polypeptide.

Monoclonal antibodies which will bind to many of these antigens are already known but in any case, with today's techniques in relation to monoclonal antibody technology, antibodies can be prepared to most antigens. The antigen-binding portion may be a part of an antibody (for example a Fab fragment) or a synthetic antibody fragment (for example a single chain Fv fragment [ScFv]). Suitable monoclonal antibodies to selected antigens may be prepared by known techniques, for example those disclosed in *"Monoclonal Antibodies: A manual of techniques"*, H Zola (CRC Press, 1988) and in *"Monoclonal Hybridoma Antibodies: Techniques and Applications"*, J G R Hurrell (CRC Press, 1982).

Chimaeric antibodies are discussed by Neuberger *et al* (1988, *8th International Biotechnology Symposium* Part 2, 792-799).

Suitably prepared non-human antibodies can be "humanized" in known ways, for example by inserting the CDR regions of mouse antibodies into the framework of human antibodies.

Further aspects of the invention provide methods (a) for determining the susceptibility of a patient to cancer comprising the steps of (i) obtaining a sample containing protein derived from the patient; and (ii) determining the relative amount or size in the said sample of the polypeptide according to the eighth aspect of the invention or determining whether there is a truncation of, or loss of function of, a polypeptide according to the eighth aspect of the invention; (b) of diagnosing cancer in a patient comprising the steps of (i) obtaining a sample containing protein derived from the patient; and (ii) determining the relative amount or size in the said sample of the polypeptide according to the eighth aspect of the invention; and (c)

of predicting the relative prospects of a particular outcome of a cancer in a patient comprising the steps of (i) obtaining a sample containing protein derived from the patient; and (ii) determining the relative amount in the said sample of the polypeptide according to the seventh aspect of the invention.

Typically, compared to a normal cell, the protein in the cancer cell is truncated or the amount of protein product is decreased.

By "derived from the patient" we include a sample directly derived from the patient or indirectly derived from, for example the protein may be produced from isolated DNA from the patient by *in vitro* transcription and translation. The sample may be any suitable sample and includes biopsy material, tumour samples (for example, those on fixed paraffin mounts and fresh and frozen tissue) and cells shed from tumour samples.

These methods are suited to determining the susceptibility of a patient to any cancer but are particularly suited to prostate cancer, melanoma, glioma or non-Hodgkin's lymphoma. Accordingly, at least for the determination of susceptibility to prostate cancer, the patient is male. Prostate cancer is particularly relevant.

Conveniently, the said polypeptide is detected using a molecule as defined in the ninth aspect of the invention. Preferably, the molecule is an antibody-like molecule comprising complementarity-determining regions specific for the polypeptide. Suitably, the molecule, such as a monoclonal antibody, comprises a detectable label. Suitable detectable labels include radioactive labels such as ^{125}I and ^{131}I and other radionuclides such as those used in diagnostic imaging, as well as any convenient fluorescent or chemiluminescent label which can readily be incorporated into the

molecule, such as an antibody. In addition the term "detectable label" also includes a moiety which can be detected by virtue of binding to another moiety (such as biotin which can be detected by binding to streptavidin); and a moiety, such as an enzyme, which can be detected by virtue of its ability to convert a colourless compound into a coloured compound, or vice versa (for example, alkaline phosphalase can convert colourless o-nitrophenylphosphate into coloured o-nitrophenol).

Conveniently, the antibodies are raised to peptides encoded by different exons of the said polypeptide. These can be used to detect truncated proteins, for example in tissue sections, as well as in protein truncation assays, and can also be used to detect changes in the level of proteins.

A further aspect of the invention provides the use of a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215 in the manufacture of a reagent for diagnosing cancer, especially prostate cancer; and in the manufacture of a medicament for treating cancer.

A still further aspect of the invention provides a method of treating cancer comprising the step of administering to the patient a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215, the nucleic acid encoding, optionally when inserted into the patient, a tumour-suppressing molecule. Tumour suppression may be identified by transfecting a (preferably prostate) tumour cell line with an expression vector comprising the polynucleotide and comparing the tumorigenic properties of the transfected cell line with the parental line in a xenograft model (eg nude mice).

Preferably, the method is for treating prostate cancer. More preferably, the nucleic acid is a tumour suppressor gene which, in this context, is a therapeutic gene. The wild-type tumour suppressor gene is preferred. Still more preferably the nucleic acid comprises a suitable delivery system.

Although adenovirus derived vectors are suited for the repair of gene defects in resting or slowly dividing tissue cells, retrovirus derived vectors specifically target rapidly dividing cells (eg tumour cells) and are therefore suited for the *in vivo* treatment of cancer therapies.

Both the amount of therapeutic protein produced and the duration of production are important issues in gene therapy. Consequently, the use of viral vectors capable of cellular gene integration (eg retroviral vectors) may be more beneficial than non-integrating alternatives (eg adenovirus derived vectors) when repeated therapy is undesirable for immunogenicity reasons.

Where the therapeutic gene is maintained extrachromosomally, the highest level of expression is likely to be achieved using viral promoters, for example, the Rous sarcoma virus long terminal repeat (Ragot *et al* (1993) *Nature* **361**, 647-650; Hyde *et al* (1993) *Nature* **362**, 250-255) and the adenovirus major late promoter. The latter has been used successfully to drive the expression of a cystic fibrosis transmembrane conductance regulator (CFTR) gene in lung epithelium (Rosenfeld *et al* (1992) *Cell* **68**, 143-155). Since these promoters function in a broad range of tissues they may not be suitable to direct cell-type-specific expression unless the delivery method can be adapted to provide the specificity. However, somatic enhancer sequences could be used to give cell-type-specific expression in an extrachromosomal setting.

Where withdrawal of the gene-vector construct is not possible, it may be necessary to add a suicide gene to the system to abort toxic reactions rapidly. The herpes simplex virus thymidine kinase gene, when transduced into cells, renders them sensitive to the drug ganciclovir, creating the option of killing the cells quickly.

The use of ectotropic viruses, which are species specific, may provide a safer alternative to the use of amphotropic viruses as vectors in gene therapy. In this approach, a human homologue of the non-human, ectotropic viral receptor is modified in such a way so as to allow recognition by the virus. The modified receptor is then delivered to cells by constructing a molecule, the front end of which is specified for the targeted cells and the tail part being the altered receptor. Following delivery of the receptor to its target, the genetically engineered ectotropic virus, carrying the therapeutic gene, can be injected and will only integrate into the targeted cells.

Virus-derived gene transfer vectors can be adapted to recognize only specific cells so it may be possible to target the cancer cell, such as prostate tumour cell. Similarly, it is possible to target expression of the therapeutic gene to the cancer cell, particularly prostate cell, using a prostate-specific promoter such as that for the PSA gene.

A further aspect of the invention provides a method of treating cancer comprising the step of administering a molecule according to the ninth aspect of the invention to the patient, the said molecule further comprising a cytotoxic moiety. The cytotoxic moiety may be directly cytotoxic (such as ricin, a suitable drug or suitable radionuclide) or it may be indirectly cytotoxic (such as an enzyme which is capable of converting a relatively non-toxic pro-drug into a relatively toxic drug; see for example WO

88/07378 and WO 91/11201).

Suitably, the molecule according to the ninth aspect of the invention is an antibody, preferably monoclonal antibody, or fragment thereof.

The aforementioned compounds of the invention or a formulation thereof may be administered by parenteral (eg subcutaneous or intramuscular) injection but preferably into the tumour. The treatment may consist of a single dose or a plurality of doses over a period of time.

Whilst it is possible for a compound of the invention to be administered alone, it is preferable to present it as a pharmaceutical formulation, together with one or more acceptable carriers. The carrier(s) must be "acceptable" in the sense of being compatible with the compound of the invention and not deleterious to the recipients thereof. Typically, the carriers will be water or saline which will be sterile and pyrogen free.

Further aspects of the invention provides for the use of a molecule according to the ninth aspect of the invention for the manufacture of a medicament for treating cancer.

It is particularly preferred that for the diagnostic methods and uses of the invention that any nucleic acid used in such methods, is a nucleic acid capable of selectively hybridising to the gene corresponding to the cDNA insert of clone IMAGE 264611.

It is particularly preferred that for the methods of treatment of the invention which use a tumour suppressor gene that the gene is the gene corresponding to the cDNA insert of clone IMAGE 264611 or a suitable variant thereof, for example a truncated version or an intron-free version

such as a cDNA. It is particularly preferred that the polypeptide capable of being encoded by a nucleic acid comprising a tumour gene which nucleic acid is capable of selectively hybridising to the said region of human chromosome 10 is a polypeptide capable of being encoded by the gene corresponding to the cDNA insert of clone IMAGE 264611.

Abbreviations used: SSCP, single-strand conformation polymorphism; PCR, polymerase chain reaction; YAC, yeast artificial chromosome; CEPH, Centre d'Etude du Polymorphisme Humain.

Brief description of the Figures and certain Tables

Figure 1 shows *a*. Examples of allele loss at microsatellite markers on 10q23-q25 in prostate tumours. The upper boxed figure beneath each peak gives the length of the allelic fragment; the lower figure is the relative peak height. 'Shoulder' peaks to the left of the main peaks are due to polymerase slippage during PCR. *b*. Microsatellite instability. Instability, thought to result from DNA mismatch repair errors (10), was seen in 1/37 tumours at 21/24 loci. Fragment lengths are given beneath each peak. The example shown here probably reflects deletion of the 207bp allele in conjunction with expansion of the 213bp allele.

Figure 2 shows allele loss at 10q23-q25. Tumour numbers correspond to those in Figure 4. Marker numbers in italics are *D10S* numbers (7). Markers denoted 'AFM' have yet to be assigned *D* numbers; the full marker names are AFMa051tb9, AFMa124wd9, AFMa064za5, AFMa301ex1 and AFMa273ye1. Tumours 8, 16, 24, 30 and 31 also show allele loss at markers *D10S189* and/or *D10S570* on the p-arm of chromosome 10, implying whole chromosome loss. The smaller numbers give the approximate genetic distance between markers in centiMorgans.

There is a clearly defined common region of deletion between markers AFMa124wd9 and *D10S583*, a distance of approximately 9 centiMorgans. By contrast, only tumours 1 and 11 show specific loss of markers around *Mxi1* and in both instances this is in conjunction with allele loss in the AFMa124wd9-*D10S583* region.

Figure 3 shows *Mxi1* loss in prostate tumours: assessment of allele loss at the $(AAAAC)_n$ polymorphism in the 3′ untranslated region of the *Mxi1* gene in tumours 1 and 11, which show specific loss of adjacent microsatellite markers, by fluorescence based typing. The boxed numbers beneath each peak give the allele fragment length (upper) and relative peak height (lower). Tumour 1 shows clear loss of *Mxi1* (peak height reduction 58%) whereas tumour 11 shows no apparent loss of *Mxi1*, despite showing loss of adjacent microsatellite marker AFMa273ye1.

Table 2 shows the results of assessment of prostate tumours for 10q23-q25 loss.

Figure 4(a) is a physical map of the minimal region showing the position of the YAC clones and markers D10S541 and AFM337, and Figure 4(b) is a more detailed map showing the position of BAC and PAC clones.

Figure 5 shows further, more informative, LOH data.

Tables 3 to 22 describe the sequenced inserts of the expressed sequence tags (ESTs) which are derived from the gene corresponding to the cDNA insert of IMAGE clone 264611.

Figure 6 (SEQ ID No 11) shows the complete sequence of a cDNA of a particularly preferred nucleic acid molecule. Potential position of introns

is shown (the "ss" above a dinucleotide denotes the splice site). The 3' untranslated sequence is in lower case.

Figure 7 (SEQ ID No 12) shows the translation in one reading frame of the nucleotide sequence of Figure 6.

Figures 8 to 15 (SEQ ID No 13 to 20) show the sequence of exons from the gene corresponding to IMAGE clone 264611 and flanking intronic sequences. Coding sequence is in upper case and intronic sequence is in lower case. PCR amplimers are in bold type. Although the exons are numbered consecutively, there may be more upstream or downstream exons and each given "exon" may be subdivided into smaller exons. R = a purine.

Example 1: Localization of a prostate tumour suppressor gene to the 10q23-q24 boundary

Materials and Methods

DNA preparation. Tumours and venous blood samples were obtained from men undergoing transurethral resection of the prostate. Tumour tissue was microdissected away from normal tissue and tumour and blood DNA prepared as described previously (6).

PCR. PCR was performed in 50μl reactions containing 30ng template DNA, 1x PCR buffer (Boehringer Mannheim), 20pmol primer, 20μM dNTPs (Boehringer Mannheim) and 1 unit of Taq polymerase (Boehringer Mannheim) on a GeneAmp 9600 thermal cycler (Perkin-Elmer Cetus). For amplification of microsatellite CA repeat markers (7) one of the primers was tagged with a fluorescent label (JOE, FAM, HEX or

TAMRA; Applied Biosystems). Microsatellite reaction mixtures were given 30 cycles of 30 seconds @ 94°, 30 seconds @ 55° and 30 seconds @ 72°, preceded by a 3 minute hot start at 95°. The annealing temperature was lowered to 50° for amplification of *Mxi1* helix-loop-helix and leucine zipper exons (5), and increased to 60° for amplification of the 3′ exon; primer sequences for 3′ exon amplification are 5′-GAGATTGAAGTGGATGTTGAAAG-3′ (SEQ ID No 7) (A) and 5′-AAATACAGGTCCTCTGACCC-3′ (SEQ ID No 8) (B) and give a 319 or 324bp product. To facilitate fluorescence based typing of the $(CAAAA)_n$ polymorphism, primer A was tagged with FAM.

Allele typing. Microsatellite allele sizes and loss of heterozygosity were determined by size separation of PCR products in a 6% denaturing polyacrylamide gel in the presence of a 2500-ROX size standard (Applied Biosystems) and detection with an 373A DNA sequencer running Genescan software (Applied Biosystems), following the manufacturer's guidelines. Up to 10 markers, distinguishable by size or fluorescent tag, were typed simultaneously. The resulting data were analysed using Genotyper software (Applied Biosystems).

It is also possible to detect LOH and to assess allele loss by staining the gel with ethidium bromide and visualizing the PCR products using a UV source, or transferring the products to a nylon or nitrocellulose membrane and hybridising with a radioactive probe derived from the marker DNA sequence (such as a radiolabelled oligonucleotide used as a primer in the initial PCR amplification). In this case the PCR products are detected by exposure of the filter to an X-ray film and allele loss may be assessed by eye or, alternatively, by densitometry.

SSCP. Following amplification of *Mxi1* introns, 10μl of PCR products

were mixed with 10μl formamide and heated to 90°C for 3 minutes. The denatured products were run in a 6% non-denaturing polyacrylamide gel at 25W for 4-6 hours with fan-assisted cooling to maintain a temperature of less than 25°C (8). DNA was transferred to a nylon membrane (Hybond N+; Amersham) and hybridized at 68°C for 3-4 hours with a mixture of both PCR primers following end labelling with 32P-dCTP (Amersham) using Terminal Transferase (Gibco-BRL). After washing in 2x SSC/0.1% SDS for 5-10 minutes, filters were exposed to X-ray film for 1-24 hours at -70°C.

DNA Sequencing. Following purification by passage through a Centricon-100 column (Amicon), PCR-amplified *Mxi1* exons were sequenced using a PRISM cycle sequencing kit (Applied Biosystems) and a 373A DNA sequencer running 373A collection and analysis software (Applied Biosystems) in accordance with the manufacturer's instructions. Each exon was sequenced twice (once from each end) from independent PCR reactions. Sequence electropherograms were aligned using Sequence Navigator software (Applied Biosystems) and compared by eye.

Results

A total of 37 prostate tumours of various and histopathological grades and stages (Table 2) were typed for loss of heterozygosity at 24 CA repeat markers spanning 10q23-q25 (7). Tumour tissue was microdissected away from normal tissue prior to DNA extraction and tumour microsatellite profiles compared to those from lymphocyte DNA to determine allele loss. 8 samples of benign hyperplastic tissue were also studied. We considered a tumour DNA sample to be showing allele loss if a reproducible signal reduction of greater than 20% as compared to normal tissue was observed, although in practice the degree of reduction was frequently much greater

and in some instances approached 100%. Examples of allele loss are shown in Figure 1. A total of 23 tumours (62%) showed allelic loss at one or more markers on 10q23-q25 (Table 2). Of these, 8 showed loss at all informative markers typed, and of these 8 a further 5 also showed allele loss at markers on the p arm, suggesting absence of the entire chromosome, possibly through non-disjunction. The allele loss data are summarized in Figure 2. No loss was seen in the benign hyperplastic tissue samples. One tumour showed microsatellite instability at the majority of loci (21/24; see Figure 1), presumably due a defective DNA mismatch correction system (10). There is no clear correlation of loss of 10q with tumour stage or grade, suggesting that 10q losses may occur at any time during tumour progression.

The retinol binding protein 4 gene (*RBP4*) and the cytochrome P450IIC gene cluster (*CYP2C*) were positioned on the deletion map following the identification of yeast artificial chromosome (YAC) clones bearing both these loci and adjacent microsatellite markers *D10S185* and *D10S571* (11). The map clearly reveals a common region of deletion proximal to *RBP4* and *CYP2C* which have been cytogenetically mapped to 10q23-24 and 10q24.1 respectively (12, 13) (Figure 2). This region is lost in all of the tumours showing 10q loss in our study, with the exception of tumour 37, which was not informative for the markers from this area. Tumours 1, 3, 6, 13, 14 and 15 define a minimal region of deletion between markers AFMa124wd9 and *D10S583*, a distance of approximately 9 centiMorgans.

Eagle *et al* have recently identified mutations in the *Mxi1* gene at 10q25 in a small number of prostate tumours, leading to speculation that Mxi1 can act as a tumour suppressor (1, 5). We were able to place *Mxi1* on the deletion map after confirming its presence on CEPH mega-YACs 936-h-5 and 966-h-9, which have been shown to overlap with YACs bearing the

microsatellite marker *D10S597* (14). Only two tumours, 1 and 11, showed specific loss of markers immediately flanking *Mxi1* and in both cases this was in conjunction with allele loss in the AFMa124wd9-*D10S583* region (Figure 2).

In an attempt to further clarify the role of *Mxi1* loss in tumour progression, we screened tumours 1 and 11, and those tumours showing loss of the entire region, for *Mxi1* mutations by PCR amplification of individual exons followed by SSCP analysis (8). Primers for PCR amplification of exons encoding helix-loop-helix and leucine zipper domains were taken from Eagle *et al* (5). For amplification of the final 3′ exon, primers derived from the immediate 5′ end of the exon and from within the 3′ untranslated sequence were used (4, 5). These 3 pairs of primers give 66% coverage of the coding sequence of *Mxi1*. The genomic structure of the 5′ end of the *Mxi1* gene has not yet been determined and we were therefore unable to analyse exons 5′ to the helix-loop-helix domain. SSCP analysis failed to detect any mutations in the two-thirds of *Mxi1* coding sequence covered.

In addition to SSCP analysis we directly sequenced those exons which encode the helix-loop-helix and leucine zipper domains previously shown to be mutated in prostate tumours (5). Again no mutations were detected. Although we were unable to detect *Mxi1* mutations in any of the tumours by either approach, we did detect a common polymorphism in the 3′ untranslated region by SSCP which subsequent sequence analysis showed to result from length variation in a $(AAAAC)_n$ tandem repeat, giving two alleles, $(AAAAC)_4$ and $(AAAAC)_5$. Eight of the tumours showing loss of the entire 10q23-q25 region or allele loss at CA repeat markers in the vicinity of *Mxi1* (Nos. 1, 8, 11, 16, 17, 21, 23 and 30) were heterozygous for this polymorphism, making it possible to assess these tumours for

actual *Mxi1* loss. 6 of the tumours (1, 8, 16, 17, 23 and 30) showing loss of adjacent markers also showed loss of *Mxi1* as determined by fluorescence based typing (Figure 3). Of these, 5 showed loss of the entire 10q24-q25 region (Figure 2). Therefore, from a total of 23 tumours showing 10q23-q25 losses, we were able to identify only one tumour (No. 1) showing specific deletion of *Mxi1* (as opposed to loss of other 10q23-q25 regions or of the entire region), and this was in conjunction with deletion of AFMa124wd9-*D10S583*.

We were also able to use this polymorphism to determine the effect of contaminating normal tissue on the efficiency of mutation detection in tumours by cycle sequencing. Exon 5 including the immediate 3′ untranslated DNA was sequenced in those tumours showing *Mxi1* loss (tumours 1, 8, 16, 17, 23 and 30). For tumour 8, which showed the greatest degree of loss of the deleted allele, the retained allele was clearly identified. The remaining tumours gave highly ambiguous sequence data following the (AAAAC)n repeat, resulting from combined termination products from the two alleles (not shown). It is therefore likely that any disabling mutations resulting from small deletion or insertion events in the retained copies of *Mxi1* would have been detected by cycle sequencing.

Discussion

The data presented here indicate the presence of a prostate tumour suppressor gene (or genes) at the 10q23-q24 boundary, and more specifically between markers AFMa124wd9 and *D10S583*, a region spanning approximately 9 centiMorgans. This region was deleted in 22 of 23 prostate tumours showing 10q losses, with the 23rd being uninformative for the relevant markers. 10q loss may be an early event in some instances of prostate carcinogenesis; losses were observed in early

as well as late stage tumours. Alternatively, 10q loss may be more important in progression of the established tumour rather than genesis given that losses were not observed in benign hyperplastic tissue samples. However, the relationship between benign prostatic hyperplasia and carcinogenesis is unclear at present and such lesions may not be a precursor to malignancy.

Although *Mxi1* has been shown to be mutated in prostate tumours, only a small proportion of cells in each tumour were found to be carrying *Mxi1* mutations (5). The authors offer two possible explanations. The first is that the tumours studied may have contained significant amounts of non-neoplastic tissue. The second is that mutated *Mxi1* alleles are only present in a small number of neoplastic cells. Given that we were unable to detect *Mxi1* mutations in microdissected tumours containing <30% contaminating normal tissue and showing a degree of 10q loss ranging from 25-79% (as estimated by microsatellite allele loss - see Table 2) the latter seems more likely. This also implies that mutation of the retained *Mxi1* allele occurs after loss of the deleted allele. The combined evidence of no mutation detection, or detection in only a small percentage of tumour cells, coupled with the allele loss data indicates the presence of a tumour suppressor gene (or genes) at 10q23-q24 of greater significance than *Mxi1* in prostate tumour progression.

Loss or rearrangement of 10q24-q25 is not restricted to prostate adenocarcinoma; it has also been observed in melanoma, glioma and non-Hodgkins lymphoma (15-21), suggesting the presence of a tumour suppressor gene or genes at this location of relevance to several tumour types.

Example 2: Identification of DNA containing a tumour suppressor gene

Figures 4 and 5 give more detailed mapping data between AFM124 and D10S583, the markers that define the minimal region in the manuscript, allowing us to narrow the minimal region further to the interval between D10S541 and D10S215; more particularly between D10S541 and AFM337xf9, a distance of less than 1cM. The physical mapping data are summarised below:

Table 1: Minimal region yeast artificial chromosome (YAC) sequence tag sites (STS) assignments

YAC	Approx size (KB)	D10S579	D10S215	AFM337xf9	D10S541
746-H-8	1200	+	+	+	+
821-D-2	1150	+	+	+	+
831-E-5	1110	+	+	+	+
921-F-8	1570	+	+	+	+
738-B-12	1330	+	+		
796-D-5	800			+	+
829-E-1	1130				+
678-F-1	480	+	+		
839-B-1	320				+
734-B-4	280				+
7B-F12	190		+	+	
24G-A10	640				+

All of these YACs other than 7B-F12 and 24G-A10 are publicly available

from the CEPH mega-YAC library. 7B-F12 and 24G-A10 are publicly available from the ICI YAC library. Both of these libraries are publicly available from the Human Genome Mapping Project Resource Centre, Hinxton Hall, Hinxton, Cambridgeshire, CB10 1RQ, UK. Sizes for mega-YAC clones are taken from CEPH data. ICI YAC clones were sized by us.

+ = STS assigned to YAC.

YACs 821-D-2, 831-E-5, 796-D-5, 24G-A-10 and 734-B-4 have been mapped in more detail to give a large scale restriction map of the region (see Figure 4). This contig does not include all restriction sites. YACs 821-D-2 and 831-E-5 appear to be identical and span the minimal region (D10S541 - AFM337xf9). They therefore contain all or part of the tumour suppressor gene.

ESTs (Expressed Sequence Tags) are generated and assigned to genomic regions using the following procedure.

1. Construct cDNA library from the tissue of interest.

2. Select individual clones at random and perform a single sequencing pass to give approximately 200-300bp of DNA sequence (an EST).

3. Design primers from each EST to allow PCR amplification of an internal fragment (an expressed Sequence Tagged Site or eSTS).

4. 'Bin' ESTs to chromosomes by PCR amplification from monochromosomal cell hybrid DNA (a panel of DNA samples derived from human/rodent cell hybrids, each of which has a single

human chromosome).

5. Localize ESTs further by PCR amplification from pools of overlapping YAC clones and finally by PCR assignment to individual YACs.

The polypeptide encoded by the cDNA insert of IMAGE clone 264611 has some similarity to the protein tensin and to auxilin, a protein involved in protein transport to the cell membrane *via* clathrin coated vesicles. The gene corresponding to the cDNA insert of clone IMAGE 264611 is a tumour suppressor gene.

The prostate tumour suppressor gene or genes are identified by screening a panel of RNAs from prostate and other tumour cell lines, in order to identify an altered, usually reduced, level of transcript. The transcript is likely to be large, as it will probably have a complex function and several sites for disabling mutation 'hits' (cf BRCA1, RB). Cross-species conservation is a good indicator that the gene has a basic cell 'housekeeping' function, the loss of which can lead to a loss of growth control and tumour formation. The prostate tumour suppressor gene cDNA is identified as follows.

Part of one of the YAC clones is used as a probe to screen a prostate cDNA library directly following radiolabelling. The 400kb *Mlu*1 fragment (marked on the restriction map in Figure 4), which covers about 75% of the minimal region, is used as a probe - this fragment can be separated cleanly from a pulsed field gel following digestion. Alternatively, the entire 24G-A10 YAC is used as a probe. A standard colony/filter hybridization approach is used. Suitable BAC or PAC clones may also be used.

Mutation analysis of the entire coding region in tumours shows that the gene is a prostate tumour suppressor gene. This is done by analysing each exon individually for mutations. Methods for mutation analysis used are single-stranded conformation polymorphism (SSCP) analysis (or variations of this technique) and direct DNA sequencing.

Genes located within the region are identified by screening of cDNA libraries with the probes obtained from the human nucleic acid sequence contained within the YACs, BACs and PACs or by exon trapping methods or by sequencing of the human nucleic acid sequence contained within the YACs, BACs and PACs, automated sequencing techniques make this routine, and use of computer programs, eg GRAIL II, that distinguish coding sequence. The results are confirmed by RT-PCR of prostate RNA from prostate tissue or a cDNA library.

The prostate tumour suppressor gene or genes are found to be expressed in normal prostate tissue, mutation analysis of the entire coding region shows that expression of the gene(s) may be altered in prostate tumours compared to normal prostate, the product of the genes may be truncated at the protein level, the mRNA product may be truncated, or have altered splicing compared to normal which results in an abnormal protein, the resulting protein encoded by the altered gene may have abnormal properties or distribution within the tissue.

Example 3: Diagnostic applications of nucleic acids

Chromosomal deletions in a specific region on chromosome 10 (ie the tumour suppressor gene-containing region at the 10q 23-q24 boundary) are detected using interphase fluorescent *in situ* hybridisation (FISH) on cells in interphase to check for loss. Cells from a biopsy sample are spread

across a slide and the cell membrane permeabilised. This allows the reagents for *in situ* hybridisation to enter the cells containing interphase chromosomes. The BACs or PACs or other suitable probes specific for the region deleted are hybridised to the chromosomes after labelling the probes with a fluorescent dye. A chromosome containing a region of deletion shows no signal; and chromosomes from a cell in which one chromosome 10 has suffered a deletion from this region will show only one signal and not two. Therefore, a method is provided that can detect 10q deletions in biopsies from patients. These are useful indicators of the staging of the grade of the tumour between benign and malignant hyperplasia and may indicate that a more aggressive treatment regime should be undertaken.

Suitable YAC clones, for use as probes, include 821-D-2, 831-E-5, 796-D-5, 24G-A-10 and 734-B-4.

Any of the BAC or PAC clones derived from the region of interest (see physical map) may be used and include 60C5 and 46B12.

It is particularly useful to use a nucleic acid which is capable of selectively hybridising to the gene corresponding to the cDNA insert of clone IMAGE 264611. The gene itself, or a suitably sized fragment thereof, is particularly suited as a probe.

The probe is ideally between 10kb and 1Mb, preferably between 60-200kb.

FISH is described by Bentz *et al* (1994) *Leukemia* 8(9), 1447-1452.

The BAC or PAC clone (such as BAC clone 60C5) is used on nuclei

isolated from prostate tissue. The method for isolating nuclei from frozen tissue is as follows.

Extraction of Nuclei from Frozen Tissue (adapted from Xiao *et al* (1995) *Am. J. Pathol.* **147**, 896-904)

(1) Cut 2x5x5 mm portion of frozen tissue - take without defrosting whole specimen. Thaw at room temperature for 1-3 minutes. (2) Mince tissue finely in 35 mm plastic petri dish using opposed scalpel blades. (3) Add 1 ml of 0.5% pepsin in 0.9% NaCl pH 1.5 to the dish. Transfer to 15 ml centrifuge tube. (4) Incubate in water bath at 37°C for 15-30 minutes or until most tumour chunks have disappeared. (NB the time taken should be the minimum required to disaggregate the tumour). Vortex every 5 minutes. (5) Add 14 ml of PBS and collect nuclei by centrifugation - 5 minutes at 15,000 rpm. (6) Discard all but 0.5 ml of supernatant by aspiration. Resuspend nuclear pellet in the residual supernatant. (7) Apply a drop (10 μl) of suspension onto a non-coated slide. Assess suspension by phase microscopy before drying to determine whether the cell density is appropriate - if nuclei are overcrowded dilute suspension with PBS; if nuclei are sparse add another drop of suspension to the same spot. (8) Air dry the slides. (9) Immerse in 10% buffered formalin for 10 minutes. (10) Air dry. (11) Bake at 55°C for 2 hours on hot plate. Slides may be stored at this point as follows (dehydrate through ethanol series (75%, 85%, 95% for 2 minutes each; air dry; store slides at -20°C with dessicant; store residual nuclear suspension in PBS at -70°C (it can be thawed x2 without any effect on the quality of the subsequent hybridization)). (12) Before hybridization the DNA needs to be denatured. Place slide on hot plate at 73°C with 70% formamide/2xSSC pH 7.0 under a coverslip for 2.5 minutes. (13) Dehydrate in ethanol series of icecold 70%, 95% and 100% for 3 minutes each and air dry.

Hybridization

Each hybridization event usually occupies half a slide. Probe labelling. The BAC or PAC clone (eg BAC clone 60C5) is used as a diagnostic probe. The whole clone is used to generate a labelled probe. A commercially available clone that recognises sequences at the centromere of chromosome 10, eg Oncor D10Z1 α-satellite, is used as a control to detect chromosome 10. The two probes are labelled differently so that they may be distinguished. The probes are to be labelled by nick translation with biotin or digoxygenin using a commercially available kit (eg Bionick kit, Life Technologies). In an Eppendorf tube mix 20ng labelled probe + 4 μg Cot 1 DNA + 2 volumes of ethanol. Dry mixture in a speed vac for 25-30 minutes. Resuspend in 11 μl hybridization mix (2xSSC, 50% formamide, 10% dextran sulphate, 1% Tween 20, pH 7.0). (If 2 or 3 probes have to be hybridized simultaneously then the 12 μl of hybridization mix should be divided equally between them (ie 2 probes 6 μl of hybridization mix each); they should not be put together until after the preannealing stage).

Denature the probe at 85°C for 5 minutes. Place immediately on ice for a few secs only. Spin quickly to get all the liquid to the bottom of the tube. Pre-anneal at 37°C for 30 minutes (after this mix 2 or more probes if necessary). Pre-annealed probe is placed on one half of a slide and covered with a 22x22 mm coverslip. Seal around coverslip with rubber solution.

Post hybridization washes

(Steps should now be carried out in the dark ie in a covered coplin jar)

3 x 5 minutes in 50% formamide, 2xSSC, pH 7.0 at 42°C

3 x 5 minutes 2xSSC, pH 7.0 at 42°C

1 x 3 minutes 4xSSC, 0.05% Tween 20, pH 7.0 (=SSCCT) at room temp

Probe detection

Step 1 - preincubation with SSCTM. Place 100 μl of SSCTM (=SSCT + 5% Marvel = 10 mls of SSCT + 0.5 g Marvel, spun down before use to remove solids) onto the slide under a 22x50 mm coverslip. Place in a moist chamber at 37°C for 10 minutes. Wash in SSCT for 3 minutes. (NB All detection reagents are diluted in SSCTM.) For each detection step 100 μl of detection reagent is placed under a 22x50 mm coverslip and placed in a moist chamber at 37°C for 25-30 minutes. Each step is followed by 3x3 minutes washes in SSCT at room temperature - the coplin jar during these steps should be shaken gently - *except* the last step which is followed by a 1 x 5 minutes wash in SSCT and 2 x 5 minutes wash in PBS. Slides are then dehydrated in an ethanol series (70%, 95%, 100% for 2 minutes each) and air dried. They are then mounted in Cytofluor (UKC ChemLab, Canterbury CT2 7NH, UK) containing DAPI 4,6-diamidino-2-phenylindole as counterstain (see below).

Dual probe detection (two colour)

Step 2 - mouse anti-Digoxygenin FITC, and Avidin-Texas Red. Step 3 - rabbit anti-mouse FITC, and anti-Avidin Biotin. Step 4 - anti-rabbit FITC, Avidin-Texas Red. Counterstain: DAPI (0.15 μg/ml = 5 μl of 30 μl/ml stock solution + 995 μl glycerol (Cytofluor).

Results

For normal prostate cells, the 60C5 probe produces two signals (spots) per cell. Two spots per cell are also seen for the chromosome 10 centromeric

marker D10Z1. If a prostate cell has only one, or no spots, produced by hybridisation with the 60C5 probe, indicating a deletion in the region covered by that probe, then the cell is cancerous. Furthermore, if the number of spots visualised using 60C5 is less than the number of spots visualised using the chromosome 10 centromeric marker, then a deletion has occurred in the region covered by 60C5, and the cell is cancerous.

The interphase FISH method can be used using genomic clones in the region. Preferably the genomic DNA is about 60-200 kb. Typically, normal tissue shows two dots, whereas tumour tissue shows one or no dots, or alternatively fewer dots than the number of chromosome 10 copies present in any cell. Centromeric repeat sequences are used to demonstrate the presence of chromosome 10 in a cell. However, even a normal tissue will show some cells with only single signals (spots). For a solid tissue, efficiency is typically between 85 and 95%, ie 85-95 nucleic per 100 show two signals. Efficiency is dependent on both the probe and the experimental conditions but may be optimised empirically. Affected tissue shows a significantly greater percentage of cells with only a single signal. The presence in the sample of contaminating, normal, cells will prevent this percentage from reaching 100%. It is therefore desirable to dissect out the area of the cells prior to these assays.

Thus, in summary, the methods and outcomes are: (i) Take tissue sample from patient, dissect out/purify affected area of tissue, and extract nuclei. (ii) Label probe with detectable tag. (iii) Contact probe with prepared sample under hybridising conditions. (iv) Remove, by washing, non-hybridised excess probe. (v) Visualise hybridised probe. Probe hybridised to a single locus is visualised as a signal (spot) by microscopy. (vi) In unaffected tissue, the majority of cells are found to show two signals, per cell. A minority of cells may show less than two spots, due

to inefficient hybridisation. (vii) In affected tissues, a significantly greater number of cells are found to show single or no signals from the specific probe. It will be appreciated that contaminating normal cells will affect the proportion of cells seen with two signals.

Prognostic information for the solid tumour, neuroblastoma, has been obtained by other workers using unrelated probes but similar FISH methods (Taylor *et al* (1994) *Br. J. Cancer* **69**, 445-451).

Example 4: Detection of polypeptides

A monoclonal antibody directed at the tumour suppressor gene product is labelled with ^{125}I. A sample of prostate tissue is prepared and proteins separated by SDS-polyacrylamide gel electrophoresis. The proteins are electroblotted onto a nitrocellulose membrane and the membrane incubated with the monoclonal antibody.

Presence of the tumour suppressor gene product is detected. The absence of the product indicates an increased susceptibility to prostate cancer.

Example 5: Therapeutic applications

The tumour suppressor gene is introduced into a patient who is susceptible to prostate cancer using a suitable retroviral vector.

Example 6: Use of IMAGE clone 264611 (and primer or probes derived therefrom) in diagnosing prostate cancer

Clone 264611 (and primers or probes derived from it) are used for detection of altered mRNA levels by *in situ* hybridisation, Northern

analysis (also detection of altered mRNA species profile) or quantitative RT-PCR. For expression detection methods (other than *in situ* hybridisation), it is preferable to use substantially pure tumour tissue. *In situ* hybridisation uses fixed tissue. A positive result indicative of prostatic cancer is altered expression levels compared to prostate tissue which is not cancerous or an altered pattern of transcript expression compared to normal prostate tissue. Samples suitable for analysis also include fresh prostate tissue, tissue collected by needle biopsy from prostate or from metastasis.

PCR primers derived from the cDNA insert of IMAGE clone 264611 are used for RT-PCR followed by mutation detection or protein truncation assays. A result indicative of prostate cancer is the detection of coding mutations, or a truncated protein product.

Thus, the methods of this Example are useful in detecting the presence of prostate adenocarcinoma.

Primers derived from intronic sequences of the gene corresponding to IMAGE 264611 (for example, those shown in Figures 29 to 34), are used to amplify the gene exons, which are then examined for mutations by various methods (sequencing, SSCP or any form of mismatch detection) or used in protein truncation assays. Suitable samples include fresh prostate tumour tissue, prostate cells recovered from blood, urine or semen, and DNA recovered from paraffin blocks.

Other methods for detecting mutations useful in this example include DGGE, direct sequencing, mis-match cleavage, heterozygote analysis and chemical cleavage.

Example 7: Loss-of-heterozygosity (LOH) as a diagnostic/prognostic tool

Loss of heterozygosity studies using markers D10S541, D10S1765 (AFM337xf9) and D10S215 are used to determine loss of the D10S541-D10S215 interval.

These markers consist of blocks of tandem CA repeats flanked by unique DNA sequence and are commonly known as microsatellites. The number of CA repeats shows variation between alleles (homologs on different chromosomes). This may be exploited to distinguish the two homologous chromosomal regions bearing these markers in a given tissue. By comparing biopsied prostate DNA (eg from urine or semen) microsatellite profiles with those of DNA extracted from blood or cheek cells (eg by means of a mouthwash), loss of one homolog of the D10S541-D10S215 interval in prostate tissue can be assessed.

This method is particularly useful for distinguishing between neoplasia (loss of one homolog) and hyperplasia (no loss) of the prostate.

The methodology for this approach is described in more detail in Example 1 and the examples given in Figure 1a and the Figure legend.

PCR primer sequences are:

D10S541:	5'-AAGCAAGTGAAGTCTTAGAACCACC-3'	(SEQ ID No 1)
	5'-CCACAAGTAACAGAAAGCCTGTCTC-3'	(SEQ ID No 2)
D10S215:	5'-TGGCATCATTCTGGGGA-3'	(SEQ ID No 3)
	5'-GCTTTACGTTTCTTCACATGGT-3'	(SEQ ID No 4)

D10S1765:	5'-ACACTTACATAGTGCTTTCTGCG-3'	(SEQ ID No 5)
	5'-CAGCCTCCCAAAGTTGC-3'	(SEQ ID No 6)

Double deletion of the gene may be detected by analogous methods.

Example 8: Mutation in tumour suppressor gene

Analysis of nucleic acid in the preferred nucleic acid of the invention, comparing a sample from a tumour with a sample from blood, revealed the following mutation:

BLOOD: GAGGCCCTAG ATTTCTATGG GGAAGT-AAGG ACCAGAGACA AAA (SEQ ID No 9)
TUMOUR: GAGGCCCTAG ATTTCTATGG GGAAGTTAAGG ACCAGAGACA AAA (SEQ ID No 10)

There is a T insertion in exon 4 (tumour 24). This mutation causes a frameshift, resulting in the incorporation of inappropriate amino-acids into the protein product following the insertion and ultimately premature truncation as the result of encountering an out-of-frame stop codon.

This mutation was detected following PCR amplification of exon 4 (using the intronic primers described in the figure of exon 4) and subsequent sequencing of the PCR product using standard methods.

TABLES

Table 2

Prostate tumours assessed for 10q23-q25 loss

TUMOUR	STAGE[a]	GRADE[b]	PATIENT AGE	10q LOSS[c]
1	T1 M0	2	81	+ (0.56)
2	T2 M0	2	84	+ (0.53)
3	T2 M0	1	67	+ (0.51)
4	T2 M0	3	70	-
5	T2 M0	3	59	-
6	T2 M0	2	64	+ (0.49)
7	T2 M1	3	84	-
8	T2 M1	3	83	+ (0.79)
9	T2 M1	3	71	-
10	T2 M1	2	83	-
11	T2 M1	2	78	+ (0.35)
12	T3 M0	3	65	IS
13	T3 M0	3	67	+ (0.65)
14	T3 M0	2	79	+ (0.46)
15	T3 M0	2	83	+ (0.52)
16	T3 M0	2	72	+ (0.36)
17	T3 M1	2	76	+ (0.37)
18	T3 M1	3	73	+ (0.60)
19	T3 M1	2	73	-
20	T3 M1	3	61	-
21	T3 M1	1	80	+ (0.57)
22	T3 M1	2	64	+ (0.34)
23	T3 M1	3	71	+ (0.25)
24	T3 M1	1	65	+ (0.56)
25	T3 M1	2	68	+ (0.38)
26	T4 M0	3	72	-
27	T4 M0	3	73	+ (0.54)
28	T4 M0	3	55	-
29	T4 M0	3	78	-
30	T4 M1	3	64	+ (0.34)
31	T4 M1	3	58	+ (0.58)
32	T4 M1	3	71	+ (0.36)
33	T4 M1	3	67	-
34	T4 M1	3	67	-
35	T4 M1	1	80	-
36	T4 M1	2	75	+ (0.62)
37	T4 M1	3	66	+ (0.38)

[a]Staging is based on digital rectal examination and bone scan (9).
[b]World Health Organization gradings: 1. Well differentiated. 2. Moderately differentiated. 3. Poorly differentiated. 4. Mixture of differentiation.
[c]+ = 10q loss - = no detected 10q loss. IS = instability. Figures in brackets give the average degree of signal reduction for microsatellite markers showing allele loss, as determined by fluorescence based typing.

LOCUS AA009519 510 bp mRNA EST 29-JUL-1996 DEFINITION ze82b09.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 365465 5' similar to SW:TENS_CHICK Q04205 TENSIN. [1] ;. ACCESSION AA009519 NID g1470718 KEYWORDS EST. SOURCE human. ORGANISM Homo sapiens Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 510) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: mob.REGA+ET High quality sequence stop: 331. FEATURES Location/Qualifiers source 1. .510 /organism="Homo sapiens" /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W." /clone="365465" /clone_lib="Soares fetal heart NbHH19W" /sex="unknown" /dev_stage="19 weeks" /lab_host="DH10B (ampicillin resistant)" mRNA <1. .>510 BASE COUNT 162 a 92 c 108 g 143 t 5 others ORIGIN AA009519 Length: 510 September 10, 1996 19:03 Type: N Check: 3385 .. 1 ATGTAGTAAG GTTTTTGGAT TCAAAGCATA AAAACCATTA CAAGATATAC 51 AATCTTTGTG CTGAAAGACA TTATGACACC GCCAAATTTA ATTGCAGAGT 101 TGCACAATAT CCTTTTGAAG ACCATAACCC ACCACAGCTA GAACTTATCA 151 AACCCTTTTG TGAAGATCTT GACCAATGGC TAAGTGAAGA TGACAATCAT 201 GTTGCAGCAA TTCACTGTAA AGCTGGAAAG GGACGAACTG GTGTAATGAT 251 ATGTGCATAT TTATTACATC GGGGCAAATT TTTAAAGGCA CAAGAGGGCC 301 CTAGATTTCT ATGGGGAAGT AAGGACCAGA GACAAAAAGG GAGTAACTAT 351 TTCCCAGTCA GAAGGCGCTA TGTGTATTAT TATTAGCTAC CTGTTAAAGA 401 ATCATCTGGA TTATAGACCA GTGGCACTGT TGTTTCCCAA GATGATGNTT 451 TGAAACTATT NCCAATGTTC AGTGGCNGGA CCTTGCAATC CNCAGTTTGT 501 GGGTCCTGCN

Table 3

LOCUS AA009520 414 bp mRNA EST 29-JUL-1996 DEFINITION ze82b09.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 365465 3'. ACCESSION AA009520 NID g1470719 KEYWORDS EST. SOURCE human. ORGANISM Homo sapiens Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 414) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: mob.REGA+ET High quality sequence stop: 317. FEATURES Location/Qualifiers source 1. .414 /organism="Homo sapiens" /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W." /clone="365465" /clone_lib="Soares fetal heart NbHH19W" /sex="unknown" /dev_stage="19 weeks" /lab_host="DH10B (ampicillin resistant)" mRNA complement(<1. .>414) BASE COUNT 104 a 71 c 72 g 165 t 2 others ORIGIN AA009520. Length: 414 September 10, 1996 19:05 Type: N Check: 5376 ..

```
1   CAGTTTATTC AAGTTTATTT TCATGGTGTT TTATCCCTCT TGATAAAAAA
51  AAATTCAGAC TTTTGTAATT TGTGTATGCT GATCTTCATC AAAAGGTTCA
101 TTCTCTGGAT CAGAGTCAGT GGTGTCAGAA TATCTATAAT GATCAGGTTC
151 ATTGTCACTA ACATCTGGTG TTACAGAAGT TGAACTGCTA GCCTCTGGAT
201 TTGACGGCTC CTCTACTGTT TTTGTGAAGT ACAGCTTCAC CTTAAAATTT
251 GGAGAAAAGT ATCGGTTGGC TTTGTCTTTA TTTGCTTTGT CAAGATCATT
301 TTTTGTTAAA GTAAGTACTA GATATTCCTT GTCATTATCT GCACGCTCTA
351 TACTGCAAAT GCTATCGATT TCTTGATCAC ATAGACTTTC CATTTTCNAC
401 TTTTTCNGAG GTTT
```

Table 4

LOCUS AA017563 241 bp mRNA EST 02-AUG-1996 DEFINITION ze39e04.s1 Soares retina N2b4HR Homo sapiens cDNA clone 361374 3'. ACCESSION AA017563 NID g1479716 KEYWORDS EST. SOURCE human. ORGANISM Homo sapiens Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 241) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Possible reversed clone: polyT not found Seq primer: -40M13 fwd. from Amersham High quality sequence stop: 166. FEATURES Location/Qualifiers source 1. .241 /organism="Homo sapiens" /note="Organ: eye; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). The retinas were obtained from a 55 year old Caucasian and total cellular poly(A)+ RNA was extracted 6 hrs after their removal. The retina RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the University of Toronto. Library constructed by Bento Soares and M.Fatima Bonaldo." /clone="361374" /clone_lib="Soares retina N2b4HR" /sex="male" /tissue_type="retina" /dev_stage="55 year old" /lab_host="DH10B (ampicillin resistant)" mRNA complement(<1. .>241) BASE COUNT 31 a 84 c 82 g 37 t 7 others ORIGIN AA017563 Length: 241 September 10, 1996 19:12 Type: N Check: 7697 .. 1 GCGGCCGCGG NGGNTGCAGC TCCANGNAGG GGGTCTGAGT CGCCTGTCAC 51 CATTTNCAGG GCTGGGAACG CCGGAGAGTT GGTCTCTCCC CTTCTACTGC 101 CTCCAACACG GCGGCNGCGG CGGCGGCACA TCCAGGGACC CGGGCCGGTT 151 TTAAACCTCC CGTCCGCCGC CGCCGCACCC CCCAGTGGCC CGGGCTCCGG 201 AGNCCGCCTG GCGGAGGCAA GCCGTTCGGA GGGATTATTC G

Table 5

LOCUS C01084 84 bp DNA EST 11-JUL-1996 DEFINITION HUMGS0007741, Human Gene Signature, 3'-directed cDNA sequence. ACCESSION C01084 NID g1433314 KEYWORDS Gene Signature; GS; EST(expressed sequence tag); BodyMap; gene expression. SOURCE One or more human adult tissue. ORGANISM Homo sapiens Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 84) AUTHORS Okubo,K. TITLE Direct Submission JOURNAL Submitted (28-DEC-1995) to the DDBJ/EMBL/GenBank databases. Kousaku Okubo, Osaka University, Institute for Molecular and Cellular Bio; 1-3,Yamada-oka, Suita, Osaka Pref. 565, Japan (E-mail:kousaku@imcb.osaka-u.ac.jp, Tel:06-877-5111(ex.3315), Fax:06-877-1922) REFERENCE 2 (bases 1 to 84) AUTHORS Okubo,K. TITLE BodyMap; human gene expression database JOURNAL Unpublished (1995) COMMENT We are not submitting the same cDNA sequence redundantly to DDBJ since 1993. For the abundance information of clones with this sequence in this library and as well as in other 3'-directed libraries, see ' http://www.imcb.osaka-u.ac.jp/bodymap'. The sequences of the clones represented by this GS sequences is also found there. FEATURES Location/Qualifiers source 1. .84 /organism="Homo sapiens" BASE COUNT 38 a 12 c 11 g 22 t 1 others ORIGIN C01084 Length: 84 September 10, 1996 19:12 Type: N Check: 5876 ..

1 GATCAGCATA CACAAATNAC AAAAGTCTGA ATTTTTTTTT ATCAAGAGGG
51 ATAAAACACC ATGAAAATAA ACTTGAATAA ACTG

Table 6

LOCUS H92038 427 bp mRNA EST 29-NOV-1995 DEFINITION ys82e12.r1 Homo sapiens cDNA clone 221326 5'. ACCESSION H92038 NID g1087616 KEYWORDS EST. SOURCE human clone=221326 primer=M13RP1 library=Soares retina N2b4HR vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). The retinas were obtained from a 55 year old Caucasian male and total cellular poly(A)+ RNA was extracted 6 hrs after their removal. The retina RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the University of Toronto. Library constructed by Bento Soares and M.Fatima Bonaldo. ORGANISM Homo sapiens Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 427) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu High quality sequence stops: 330 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. FEATURES Location/Qualifiers source 1. .427 /organism="Homo sapiens" /clone="221326" mRNA <1. .>427 BASE COUNT 103 a 75 c 116 g 129 t 4 others ORIGIN H92038 Length: 427 September 10, 1996 19:06 Type: N Check: 6168 .. 1 GGAAGTNGGT NATGGTCTTC AAAAGGATAT TGTGCAACTC TGCAATTAAA 51 TTTGGCGGTG TCATAATGTC TTTCAGCACA AAGATTGTAT ATCTTGTAAT 101 GGTTTTTATG CTTTGAATCC AAAAACCTTA CTACATCATC AATATTGTTC 151 CTGTATACGC CTTCAAGTCT TTCTGCAGGA AATCCCATAG CAATAATGTT 201 TGGATAAATA TAGGTCAAGT CTAAGTCGAA TCCATCCTCT TGATATCTCC 251 TTTTGTTTCT GGCTAACGAT CTCTTTGGAT GGATGGCTGT CATGTCTGGG 301 GAGCCTGTGN TGGNAAGGAA AAAGGGAGGG AGAGAGATGG GCAGAAGCTG 351 GCTCGGTGGG CGGGGGCTTT CTTCTGGCAG GGATGGGAAA TGGGCTCTGG 401 GGACTGGGCG GTACTGGATG GCCCCTC

Table 7

LOCUS H92039 117 bp mRNA EST 29-NOV-1995 DEFINITION ys82e12.s1 Homo sapiens cDNA clone 221326 3'. ACCESSION H92039 NID g1087617 KEYWORDS EST. SOURCE human clone=221326 primer=Promega -21m13 library=Soares retina N2b4HR vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). The retinas were obtained from a 55 year old Caucasian male and total cellular poly(A)+ RNA was extracted 6 hrs after their removal. The retina RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the University of Toronto. Library constructed by Bento Soares and M.Fatima Bonaldo. ORGANISM Homo sapiens Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 117) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu High quality sequence stops: 104 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Possible reversed clone: polyT not found. FEATURES Location/Qualifiers source 1. .117 /organism="Homo sapiens" /clone="221326" mRNA <1. .>117 BASE COUNT 16 a 44 c 37 g 19 t 1 others ORIGIN H92039 Length: 117 September 10, 1996 19:12 Type: N Check: 5577 .. 1 TCCAGGGCTG GGAACGCCGG AGAGTTGGTC TCTCCCCTTC TACTGCCTCN 51 AACACGGCGG CGGCGGCGGC GGCACATCCA GGGACCCGGG CCGGTTTTAA 101 ACCTCCCGTC CGCCGCC

Table 8

LOCUS N20238 322 bp mRNA EST 18-DEC-1995 DEFINITION yx44f06.s1 Homo sapiens cDNA clone 264611 3'. ACCESSION N20238 NID g1125193 KEYWORDS EST. SOURCE human clone=264611 primer=m13 -40 forward library=Soares melanocyte 2NbHM vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI Male. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino. ORGANISM Homo sapiens Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 322) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu High quality sequence stops: 209 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Possible reversed clone: polyT not found. FEATURES Location/Qualifiers source 1. .322 /organism="Homo sapiens" /clone="264611" mRNA <1. .>322 BASE COUNT 49 a 112 c 98 g 57 t 6 others ORIGIN N20238 Length: 322 September 10, 1996 19:07 Type: N Check: 7249 .. 1 GGTCTGAGTC GCCTGTCACC ATTTCCAGGG CTGGGAACGC NGGAGAGTTG 51 GTCTCTCCCC TTCTACTGCC TCCAACACGG CGGCGGCGGC GGCGGCACAT 101 CCAGGGACCC GGGCCGGTTT TAAACCTCCC GTCCGCCGCC GCCGCACCCC 151 CCGTGGCCCG GGCTCCGGAG GCCGCCGGCG GAGNAAGCCG TTTCGGAGGA 201 TTATTCGTCT TCTCCCCATT CCGCTGCCGC CCGCTGCCAG GCTCTTGGTG 251 CTTGAAGAAG AAGCAGGCCA GTTGNCTGAA ACCATTCNAG AAGCCGCNGA 301 AGCAGCCATT ACNCGGCTGC GG

Table 9

LOCUS N29304 427 bp mRNA EST 04-JAN-1996 DEFINITION yx44f06.r1 Homo sapiens cDNA clone 264611 5'. ACCESSION N29304 NID g1147540 KEYWORDS EST. SOURCE human clone=264611 primer=T7 library=Soares melanocyte 2NbHM vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI Male. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino. ORGANISM Homo sapiens Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates: Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 427) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu High quality sequence stops: 370 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. FEATURES Location/Qualifiers source 1. .427 /organism="Homo sapiens" /clone="264611" mRNA <1. .>427 BASE COUNT 116 a 90 c 79 g 140 t 2 others ORIGIN N29304 Length: 427 September 10, 1996 19:04 Type: N Check: 9508 .. 1 TAAGTACTAG ATATTCCTTG TCATTATCTG CACGCTCTAT ACTGCAAATG 51 CTATCGATTT CTTGATCACA TAGACTTCCA TTTTCTACTT TTTCTGAGGT 101 TTCCTCTGGT CCTGGTATGA AGAATGTATT TACCCAAAAG TGAAACATTT 151 TGTCCTTTTT TAGCATCTTG TTCTGTTTGT GGAAGAACTC TACTTTGATA 201 TCACCACACA CAGGTAACGG CTGAGGGAAC TCAAAGTACA TGAACTTGTC 251 TTCCCGTCGT GTGGGTCCTG AATTGGAGGA ATATATCTTC ACCTTTAGCT 301 GGCAGACCAC AAACTGNAGG ATTGCAAGTT CCGCCACTGA ACATTGGAAT 351 AGTTTCAAAC ATCATCTTGT GAAACAACAG TGCCACTGGT CTATAANCCA 401 GATGATTCTT TAACAGGGTA GCTATAA

Table 10

LOCUS N35389 437 bp mRNA EST 16-JAN-1996 DEFINITION yy23e03.s1 Homo sapiens cDNA clone 272092 3'. ACCESSION N35389 NID g1156531 KEYWORDS EST. SOURCE human clone=272092 primer=m13 -40 forward library=Soares melanocyte 2NbHM vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI Male. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino. ORGANISM Homo sapiens Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 437) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu High quality sequence stops: 311 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. FEATURES Location/Qualifiers source 1. .437 /organism="Homo sapiens" /clone="272092" mRNA <1. .>437 BASE COUNT 108 a 79 c 78 g 166 t 6 others ORIGIN N35389 Length: 437 September 10, 1996 19:04 Type: N Check: 9803 .. 1 CAGTTTATTC AAGTTTATTT TCATGGTGTT TTATCCCTCT TGATAAAAAA 51 AAATTCAGAC TTTTGTAATT TGTGTATGCT GATCTTCATC AAAAGGTTCA 101 TTCTCTGGAT CAGAGTCAGT GGTGTCAGAA TATCTATAAT GATCAGGTTC 151 ATTGTCACTA ACATCTGGTG TTACAGAAGT TGAACTGCTA GCCTCTGGAT 201 TTGACGGCTC CTCTACTGTT TTNGTGAAGT ACAGCTTCAC CTTAAAATTT 251 GGAGAAAAGT ATCGGTTGGC TTTGTCTTTA TTTGCNTTGT CAAGATCATT 301 TTCTGTTAAA GTAAGTACTA TGATATTCCT TGTCATTATC TGCACGCTCT 351 ATACTGCAAA TGCTATCGAT TTCTTGATCA CATAGACTTC CATTTTCTAC 401 TTTTTCNGAG GTTTCCCCCN GGTCCNGGGT AATGAAN

Table 11

LOCUS N48030 372 bp mRNA EST 14-FEB-1996 DEFINITION yy23e03.r1 Homo sapiens cDNA clone 272092 5' similar to SW:TENS_CHICK Q04205 TENSIN. [1] ;. ACCESSION N48030 NID g1189196 KEYWORDS EST. SOURCE human clone=272092 primer=T7 library=Soares melanocyte 2NbHM vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI Male. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino. ORGANISM Homo sapiens Eukaryotae; Metazoa; Eumetazoa; Bilateria: Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 372) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu High quality sequence stops: 240 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. FEATURES Location/Qualifiers source 1. .372 /organism="Homo sapiens" /clone="272092" mRNA <1. .>372 BASE COUNT 122 a 67 c 76 g 101 t 6 others ORIGIN N48030 Length: 372 September 10, 1996 19:06 Type: N Check: 6071 ..

1 TTTTTGGATT CAAAGCATAA AAACCATTAC AAGATATTTT ATCTTCTNNG

51 CTGAAAGACA TTATGACACC GCCAAATTTA ATTGCAGAGT TGCACAATAT

101 CCTTTTGAAG ACCATAACCC ACCACAGCTA GAACTTATCA AACCCTTTTG

151 TGAAGATCTT GACCAATGGC TAAGTGAAGA TGACAATCAT GTTGCAGCAA

201 TTCACTGTAA AGCTGGAAAG GGACGAACTG GTGTAATGAT ATGTGCATAT

251 TTATTACATC GGGGCAAATT TTTAAAGGCA CAAGAGGCCC NAAGATTTCT

301 ATGGGGAAGT AAGGGCCCGA GACNAAAAGG GNGTAACTAT TCCCAGTCAG

351 AGGGCGCTAT GTGTNTTATT AT

Table 12

LOCUS R06763 474 bp mRNA EST 03-APR-1995 DEFINITION yf11e03.s1 Homo sapiens cDNA clone 126556 3'. ACCESSION R06763 NID g757383 KEYWORDS EST. SOURCE human clone=126556 library=Soares fetal liver spleen 1NFLS vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) primer=SP6 Rsite1=Pac I Rsite2=Eco RI Liver and spleen from a 20 week-post conception male fetus. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo. ORGANISM Homo sapiens Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 474) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu High qality sequence stops: 108 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. FEATURES Location/Qualifiers source 1. .474 /organism="Homo sapiens" /clone="126556" BASE COUNT 108 a 81 c 89 g 190 t 6 others ORIGIN R06763 Length: 474 September 10, 1996 19:04 Type: N Check: 6789 .. 1 AGCCGCTTTA ATTAAAGATC TTTTTTTTTT TTTTTTTTTC AGTTTATTCA 51 AGTTTATTTT CATGGTGTTT TATCCCTCTT GATAAAAAAA AATTCAGACT 101 TTTGTAATTT GTGTATGCTG ATCTTCATCA AAAGGGTTCA TTCTCTGGAT 151 CAGAGTCAGT GGGTGTCAGA ATATCTATAA TGATCAGGTT CATTGTCACT 201 AACATCTGGN GTTACAGAAG TTGAACTGCT AGCCTCTGGG ATTTGACGGC 251 TCCNCTACTG TTTTTGTGAA GTACAGCTTC ACCTTAAAAT TTGGNGAAAA 301 GTATCGGTTG GCTTTGTCTT TATTTGCTTT GTCAAGATCA TTTTTTGTTA 351 AAGTAAGGAC TAGGATATTC CCTGTCATTA TCTGCACGCT CTATACTGCA 401 AATGCTATCG ATTTCTTGAT CACATAGGGC TTCCNTTTTC TACTTTTTCT 451 GAGGGTTNCC CTGGTCCGGG NTTG

Table 13

LOCUS R06814 429 bp mRNA EST 03-APR-1995 DEFINITION yf11e03.r1 Homo sapiens cDNA clone 126556 5'. ACCESSION R06814 NID g757434 KEYWORDS EST. SOURCE human clone=126556 library=Soares fetal liver spleen 1NFLS vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) primer=M13RP1 Rsite1=Pac I Rsite2=Eco RI Liver and spleen from a 20 week-post conception male fetus. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo. ORGANISM Homo sapiens Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 429) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu High qality sequence stops: 307 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. FEATURES Location/Qualifiers source 1. .429 /organism="Homo sapiens" /clone="126556" BASE COUNT 114 a 73 c 65 g 176 t 1 others ORIGIN R06814 Length: 429 September 10, 1996 19:16 Type: N Check: 889 .. 1 TGTTCTGTAA GTTACTTTTA CCGTTAAACT TCTTAATGTT GCTTATTGTT 51 TGTCTTACAT TTTTAGGTTG GATTTTTCTT AAGTCACATG TCTAATAAAA 101 AAAACCCTTA AATACCTCAT TTATTCGTCT TCGTTAGTGA ATGCATTGTT 151 GTACATATTA GATTTTTCTC TTTAGATAAC TCAGCTTCCC CTATTAAGTG 201 CCACATGTAT TACAAAATTT TATTTATGTT TTATTGTTTA ATAAACTCTT 251 GAGAACTAGA TACATTTTAA TCATTTGTAA TACTTACATT TTCTAAAACA 301 CTTCATTTTT CCCGGGGTTC TTCAACAAAG GGGATGGCAT GTAGGTACAA 351 GGGATAGCTT TACCNGTGTT AGGAAGGTTG TTTTCACACC TTTACATCAA 401 CTGCATAGTC CCGTTTTTGT TGGGGCCCA

Table 14

LOCUS R29457 224 bp mRNA EST 25-APR-1995 DEFINITION F1-578D 22 week old human fetal liver cDNA library Homo sapiens cDNA clone F1-578D 5'. ACCESSION R29457 NID g1511865 KEYWORDS EST. SOURCE human. ORGANISM Homo sapiens Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 224) AUTHORS Choi,S.S., Yun,J.W., Choi,E.K., Cho,Y.G., Sung,Y.C. and Shin,H.-S. TITLE Construction of a gene expression profile of a human fetal liver by single-pass cDNA sequencing JOURNAL Unpublished (1995) COMMENT Contact: Hee-Sup Shin Developmental Genetics Pohang Institute of Science & Technology San31, Hyojadong Pohang, 790-784 Republic of Korea Tel: 562-279-2291 Fax: 562-279-2199 Email: shinhs@vision.postech.ac.kr Seq primer: T3 primer. FEATURES Location/Qualifiers source 1. .224 /organism="Homo sapiens" /note="Vector: pBluescriptII SK(-); Site_1: EcoRI; Site_2: XhoI; The cDNA library made by oligo-dT primed and directionally cloned between 5'ExoR I-XhoI3' sites." /clone="F1-578D" /clone_lib="22 week old human fetal liver cDNA library" /lab_host="XL1-blue MRF'" mRNA <1. .>224 BASE COUNT 45 a 78 c 67 g 34 t ORIGIN R29457 Length: 224 September 10, 1996 19:11 Type: N Check: 1046 .. 1 GGGCTCCGGA GCCGCCGGCG GAGGCAGCCG TTCGGAGGAT TATTCGTCTT 51 CTCCCCATTC CGCTGCCGCC GCTGCCAGGC CTCTGCTGCT GAGGAGAAGC 101 AGGCCCAGTC GCTGCAACCA TCCAGCAGCC GCCGCAGCAG CCATTACCCG 151 GCTGCGGTCC AGAGCCAAGA CGCAGAGAGG GCATCAGCTA CCGCCAAGTC 201 AGAGCATTTC CATCTCAGAA GAAG

Table 15

LOCUS T05157 266 bp mRNA EST 30-JUN-1993 DEFINITION EST03045 Homo sapiens cDNA clone HFBCS42. ACCESSION T05157 NID g316309 KEYWORDS EST. SOURCE Human clone=HFBCS42 library=Fetal brain, Stratagene (cat#936206) vector=LambdaZAP-II primer=M13-21 17-18 wk gestation, female; oligo-dT + random primed cDNA synthesis; lambdaZAP-II vector, 1.0kb average inser size. ORGANISM Homo sapiens Eukaryota; Animalia; Chordata; Verebrata; Mammalia; Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominidae. REFERENCE 1 (bases 1 to 266) AUTHORS Adams,M.D., Kerlavage,A.R., Fields,C. and Venter,J.C. TITLE 3400 Expressed Sequence Tags Identify Diversity of Transcripts from Human Brain JOURNAL Nature Genet. 4, 256-267 (1993) COMMENT Contact: Adams, MD The Institute for Genomic Research 932 Clopper Road, Gaithersburg, MD 20878 Tel: 3018699056 Fax: 3018699423 Email: mdadams@tigr.org. FEATURES Location/Qualifiers source 1. .266 /organism="Homo sapiens" /clone="HFBCS42" BASE COUNT 95 a 44 c 57 g 69 t 1 others ORIGIN T05157 Length: 266 September 10, 1996 19:06 Type: N Check: 4398 .. 1 TGGAGGGAAG ACAAGTTCAT GTACTTTGAG TTCCCTCAGC CGTTACCTGT 51 GTGTGGTGAT ATCAAAGTAG AGTTCTTCCA CAAACAGAAC AAGATGCTAA 101 AAAAGGACAA AATGTTTCAC TTTTGGGTAA ATACATTCTT CATACCAGGA 151 CCAGAGGAAA CCTCAGAAAA AGTAGAAAAT GGAAGTCTAT GTGATCAAGN 201 AATCGATAGC ATTTGCAGTA TAGAGCGTGC AGATAATGAC AAGGAATATC 251 TAGTACTTAC TTTAAC

Table 16

LOCUS T60214 396 bp mRNA EST 09-FEB-1995 DEFINITION yc22c07.r1 Homo sapiens cDNA clone 81420 5'. ACCESSION T60214 NID g662051 KEYWORDS EST. SOURCE human clone=81420 library=Stratagene lung (#937210) vector=pBluescript SK- host=SOLR cells (kanamycin resistant) primer=M13RP1 Rsite1=EcoRI Rsite2=XhoI Normal lung tissue from a 72 year old male. Cloned unidirectionally. Primer: Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence: 5'-GAATTCGGCACGAG-3'; 3' adaptor sequence: 5'-CTCGAGTTTTTTTTTTTTTTTT-3'. ORGANISM Homo sapiens Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 396) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu High qality sequence stops: 242 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. FEATURES Location/Qualifiers source 1. .396 /organism="Homo sapiens" /clone="81420" BASE COUNT 119 a 75 c 74 g 126 t 2 others ORIGIN T60214 Length: 396 September 10, 1996 19:07 Type: N Check: 5134 .. 1 TCAAATCCAG AGGCTAGCAG TTCAACTTCT GTAACACCAG ATGTTAGTGA 51 CAATGAACCT GATCATTATA GATATTCTGA CACCACTGAC TCTGATCCAG 101 AGAATGAACC TTTTGATGAA GATCAGCATA CACAAATTAC AAAAGTCTGA 151 ATTTTTTTTT ATCAAGAGGG ATAAAACACC ATGAAAATAA ACTTGAATAA 201 ACTGAAAATG GGACCTTTTT TTTTTTTAAT GGGCAATAGG GACATTGTGT 251 CAGGATTACC AGTTATAGGG GACAATTCTC TTTTCCCTGG ACCCAATCTT 301 GTTTTTTACC CTATACATCC ACCGGGGGTT TTTTGACACT TGTTTGTCCC 351 AGTTGGAAAA AGGGTTGTNT TGGCCGTNGT CCAGGATTAT ACCCTT

Table 17

LOCUS W23656 451 bp mRNA EST 06-MAY-1996 DEFINITION zb46c05.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 306632 5'. ACCESSION W23656 NID g1300471 KEYWORDS EST. SOURCE human. ORGANISM Homo sapiens Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 451) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: mob.REGA+ET High quality sequence stop: 240. FEATURES Location/Qualifiers source 1. .451 /organism="Homo sapiens" /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCAATTTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal heart library, Soares fetal heart NbHH19W." /clone="306632" /clone_lib="Soares fetal lung NbHL19W" /dev_stage="19 weeks" /lab_host="DH10B (ampicillin resistant)" mRNA <1. .>451 BASE COUNT 148 a 76 c 82 g 141 t 4 others ORIGIN W23656 Length: 451 September 10, 1996 19:10 Type: N Check: 6961 .. 1 CAACTTCTGT AACACCAGAT GTTAGTGACA ATGAACCTGA TCATTATAGA 51 TATTCTGACA CCACTGACTC TGATCCAGAG AATGAACCTT TTGATGAAGA 101 TCAGCATACA CAAATTACAA AAGTCTGAAT TTTTTTTTAT CAAGAGGGAT 151 AAAACACCAT GAAAATAAAC TTGAATAAAC TGAAAATGGA CCTTTTTTTT 201 TTTAATGGCA ATAGGACATT GTGTCAGATT ACCAGTTATA GGAACAATTC 251 TCTTTTCCTG ACCAATCTTG NTTTACCCNA TACATTCCCA GGGGTTTGGA 301 CACTTGGTGG TCCAGNTTGA AAAAAGGTTG TGTAGCTGTG NCATGGTATA 351 TACCTTTTTG TGGCCAAAAG GGACATTTAA AATTCAATTA GGATTAATAA 401 AGATGGGCAC TTTCCCGTTT AATTCCAGTT TTATAAAAAG TGGGGACAGA 451 C

Table 18

LOCUS W27533 902 bp mRNA EST 08-MAY-1996 DEFINITION 32b2 Human retina cDNA randomly primed sublibrary Homo sapiens cDNA. ACCESSION W27533 NID g1307337 KEYWORDS EST. SOURCE human. ORGANISM Homo sapiens Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 902) AUTHORS Macke, J., Smallwood, P. and Nathans, J. TITLE Adult Human Retina cDNA JOURNAL Unpublished (1996) COMMENT Contact: Dr. Jeremy Nathans Dr. Jeremy Nathans, Dept. of Molecular Biology and Genetics Johns Hopkins School of Medicine 725 North Wolfe Street, Baltimore, MD 21205 Tel: 410 955 4678 Fax: 410 614 0827 Email: jeremy_nathans@qmail.bs.jhu.edu Clones from this library are NOT available. PCR PRimers FORWARD: CTTTTGAGCAAGTTCAGCCTGGTTAAGT BACKWARD: GAGGTGGCTTATGAGTATTTCTTCCAGGGTAA Seq primer: GGGTAAAAAGCAAAAGAATT. FEATURES Location/Qualifiers source 1. .902 /organism="Homo sapiens" /note="Organ: eye; Vector: lambda gt10; Site_1: EcoRI; Site_2: EcoRI; The library used for sequencing was a sublibrary derived from a human retina cDNA library. Inserts from retina cDNA library DNA were isolated, randomly primed, PCR amplified, size-selected, and cloned into lambda gt10. Individual plaques were arrayed and used as templates for PCR amplification, and these PCR products were used for sequencing." /clone_lib="Human retina cDNA randomly primed sublibrary" /sex="mixed (males and females)" /tissue_type="retina" /dev_stage="adult" /lab_host="E. coli strain K802" mRNA <1. .>902 BASE COUNT 124 a 110 c 117 g 131 t 420 others ORIGIN W27533 Length: 902 September 10, 1996 19:05 Type: N Check: 224 ..

1 GNGNNNTTNC TACTCANGAT CATTTGGNGG TTAAAGTAAG TACTAGATAN
51 TCCTTGTCAT TATCTGCACG CTCTATACTG CAAATGCTAT CGATTTCTTG
101 ATCACATAGA CTTCCATTTT CTACTTTTNC TGAGGTTNCC TCTGGTCCTG
151 GTATGAAGAA TGTATTTACC CAAAAGTGAA ACATTGGGTC CTTTTTTAGC
201 ATCTGGTNCT GTGNGTGGAA GAACTCTACT TGGATATCAC CACACACAGG
251 TAACGGCTGA GGGAACTCAA AGTACATGAA CTTGTCTTCC CGNCGNGTGG
301 GTCCTGAATT GGAGGAATAT NTCTTCACCT NNAGCTGGCA GACCACAAAC
351 TGAGGATTGC AAGTNCCGCC ACTGAACATG GGAATAGGNT CAAACATCAN
401 CTTGGGAAAC AACAGGGNCA CTGGTCTTTT ANCCAGNTGA TCNNNACAGG
451 GGGTATNATA NACANANGGG CCCNNNNNGG AATGGGNCNC CNNGGGGTTN
501 NNCCCNNNNC CCANNNNNNC ANNGGGNTNC CGGNGGGNNN NNNNNNNNNN
551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN

Table 19

NNNNNNNNNN NNNNNNNNNN 701 NNNNNNNNNN NNNNNNNNNN
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN 751 NNNNNNNNNN
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN 801
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
NNNNNNNNNN 901 CC

Table 19 (continued)

LOCUS W30684 601 bp mRNA EST 09-MAY-1996 DEFINITION zb77b11.r1 Soares senescent fibroblasts NbHSF Homo sapiens cDNA clone 309597 5'. ACCESSION W30684 NID g1311870 KEYWORDS EST. SOURCE human. ORGANISM Homo sapiens Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 601) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: mob.REGA+ET High quality sequence stop: 463. FEATURES Location/Qualifiers source 1. .601 /organism="Homo sapiens" /note="Vector: pT7T3D (Pharmacia) with a modified polylinker V_TYPE: phagemid; Site_1: Not I; Site_2: Eco R I ; TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo." /clone="309597" /clone_lib="Soares senescent fibroblasts NbHSF" /lab_host="DH10B (ampicillin resistant)" mRNA <1. .>601 BASE COUNT 176 a 105 c 122 g 197 t 1 others ORIGIN W30684 Length: 601 September 10, 1996 19:13 Type: N Check: 2320 ..

```
1 GCAAGAGGGA TAAAACACCA TGAAAATAAA CTTGAATAAA CTGAAAATGG
51 ACCCTTTTTT TTTTAATGGC AATAGGACAT TGTGTCAGAT TACCAGTTAT
101 AGGAACAATT CTCTTTTCCT GACCAATCTT GTTTTACCCT ATACATCCAC
151 AGGGTTTTGA CACTTGTTGT CCAGTTGAAA AAAGGTTGTG TAGCTGTGTC
201 ATGTATATAC CTTTTTGTGT CAAAAGGACA TTTAAAATTC AATTAGGATT
251 AATAAAGATG GCACTTTCCC GTTTTATTCC AGTTTTATAA AAAGTGGAGA
301 CAGACTGATG TGTATACGTA GGAATTTTTT CCTTTTGTGT TCTGTCACCA
351 ACTGAAGTGG CTAAAGAGCT TTGTGATATA CTGGTTCACA TCCTACCCCT
401 TTGCACTTGT GGCAACAGAT AAGTTTGCAG TTGGGCTAAG AGAGGTTTCC
451 GAAGGGTTTT GCTACATTCT AATGCATGTA TTCGGGGTTA GGGGAATGGA
501 GGGGAATGCT CAGAAAGGAA ATAATTTTAA TGCTGGACTC TGGACCATAT
551 ACCATCTCCA GCTANTTACA CACACCTTTC CTTAGCATGC CACAGTTATT
601 A
```

Table 20

LOCUS W81026 453 bp mRNA EST 26-JUN-1996 DEFINITION zd84a07.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 347316 5'. ACCESSION W81026 NID g1392060 KEYWORDS EST. SOURCE human. ORGANISM Homo sapiens Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 453) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: mob.REGA+ET High quality sequence stop: 392. FEATURES Location/Qualifiers source 1..453 /organism="Homo sapiens" /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W." /clone="347316" /clone_lib="Soares fetal heart NbHH19W" /sex="unknown" /dev_stage="19 weeks" /lab_host="DH10B (ampicillin resistant)" mRNA <1..>453 BASE COUNT 190 a 77 c 79 g 106 t 1 others ORIGIN W81026 Length: 453 September 10, 1996 19:03 Type: N Check: 2953 ..

```
   1 ATACCAGGAC CAGAGGAAAC CTCAGAAAAA GTAGAAAATG GAAGTCTATG
  51 TGATCAAGAA ATCGATAGCA TTTGCAGTAT AGAGCGTGCA GATAATGACA
 101 AGGAATATCT AGTACTTACT TTAACAAAAA ATGATCTTGA CAAAGCAAAT
 151 AAAGACAAAG CCAACCGATA CTTTTCTCCA AATTTTAAGG TGAAGCTGTA
 201 CTTCACAAAA ACAGTAGAGG AGCCGTCAAA TCCAGAGGCT AGCAGTTCAA
 251 CTTCTGTAAC ACCAGATGTT ACGTGACAAT GAACCTGATC ATTATAGATA
 301 TTCTGACACC ACTGACTCTG ATCCAGAGAA TGAACCTTTT GATGAAGATC
 351 AGCATACACA AATTACAAAA GTCTGAATTT TTTTTTATCA AGAGGGATAA
 401 AACACCATGG AAAATAAACT TGGAATAAAC TGAAAAANAA AAAAAAAAAA
 451 GAT
```

Table 21

LOCUS W81062 429 bp mRNA EST 26-JUN-1996 DEFINITION zd84a07.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 347316 3'. ACCESSION W81062 NID g1392114 KEYWORDS EST. SOURCE human. ORGANISM Homo sapiens Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 429) AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M.. Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R. TITLE The WashU-Merck EST Project JOURNAL Unpublished (1995) COMMENT Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: mob.REGA+ET High quality sequence stop: 324. FEATURES Location/Qualifiers source 1. .429 /organism="Homo sapiens" /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W." /clone="347316" /clone_lib="Soares fetal heart NbHH19W" /sex="unknown" /dev_stage="19 weeks" /lab_host="DH10B (ampicillin resistant)" mRNA complement(<1. .>429) BASE COUNT 105 a 83 c 77 g 161 t 3 others ORIGIN W81062 Length: 429 September 10, 1996 19:05 Type: N Check: 7359 .. 1 CAGTTTATTC AAGTTTATTT TCATGGTGTT TTATCCCTCT TGATAAAAAA 51 AAATTCAGAC TTTTGTAATT TGTGTATGCT GATCTTCATC AAAAGGTTCA 101 TTCTCTGGAT CAGAGTCAGT GGTGTCAGAA TATCTATAAT GATCAGGTTC 151 ATTGTCACTA ACATCTGGTG TTACAGAAGT TGAACTGCTA GCCTCTGGAT 201 TTGACGGCTC CTCTACTGTT TTTGTGAAGT ACAGCTTCAC CTTAAAATTT 251 GGAGAAAAGT ATCGGTTGGC TTTGTCTTTA TTTGCTTTGT CAAGATCATT 301 TTTTGTTAAA GTAAGTACTA AGATATTCCT TGTCATTATC TGCACGCTCT 351 AATACTGCAA ATGGCTATCC GATTTCCTGG ATCCACCATA GGNCTTCCNA 401 TTTCCAACTT TTCCCTGNGG TTCCCCCGG

Table 22

REFERENCES

1. Cannon-Albright, L., and Eeles, R. (1995) "Progress in prostate cancer" *Nature Genet.* 9, 336-338.

2. Lundgren, R., Mandahl, N., Heim, S., Limon, J., Henrikson, H., and Mitelman, F. (1992) "Cytogenetic analysis of 57 primary prostatic adenocarcinomas" *Genes Chrom. Cancer* 4, 16-24.

3. Arps, S., Rodewald, A., Schmalenberger, B., Carl, P., Bressel, M., and Kastendieck, H. (1993) "Cytogenetic survey of 32 cancers of the prostate" *Cancer Genet. Cytogenet.* 66, 93-99.

4. Zervos, A.S., Gyuris, J., and Brent, R. (1993) "Mxi1, a protein that specifically interacts with Max to bind Myc-Max recognition sites" *Cell* 72, 223-232.

5. Eagle, L.R., Yin, X., Brothman, A.R., Williams, B.J., Atkin, N.B., and Prochownick, E.V. (1995) "Mutation of the Mxi1 gene in prostate cancer" *Nature Genet.* 9, 249-255.

6. Phillips, S.M.A., Morton, D.G., Lee, S.J., Wallace, D.M.A., and Neoptolemos, J.P. (1994) "Loss of heterozygosity of the retinoblastoma and adenomatous polyposis susceptibility gene loci and in chromosomes 10p, 10q, and 16q in human prostate cancer" *Br. J. Urol.* 73, 390-395.

7. Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernadi, G., Lathrop, M., and Weissenbach, J. (1994) "The 1993-1994 Genethon human genetic linkage map" *Nature Genetics* 7, 246-339.

8. Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., and Sekiya, T. (1989) "Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms" *Proc. Natl. Acad. Sci. USA* 86, 2766-2770.

9. UICC (Union Internationale Contre Le Cancer). TNM

Classification of malignant tumours. Geneva: International Union Against Cancer, 1978.

10. Parsons, R., Li, G.-M., Longley, M.J., Fang, W., Papadopoulos, N., Jen, J., de la Chapelle, A., Kinzler, K.W., Vogelstein, B., and Modrich, P. (1993) "Hypermutability and mismatch repair deficiency in RER+ tumour cells" *Cell* **75**, 1227-1236.

11. Gray, I.C., Nobile, C., Moresu, R., Ford, S., and Spurr, N.K. (1995) "A 2.4 megabase physical map spanning the *CYP2C* gene cluster on chromosome 10q24" *Genomics* **28**, 328-332.

12. Rocchi, M., Covone, A., Romeo, G., Faraonio, R., and Colantuoni, V. (1989) "Regional mapping of RBP4 to 10q23-24 and RBP1 to 3q21-22 in man" *Somat. Cell Molec. Genet.* **15**, 185-190.

13. Inoue, K., Inazawa, J., Suzuki, Y., Shimada, T., Yamazaki, H., Guengerich, F.P., and Abe, T. (1994) "Fluorescence in-situ hybridization analysis and chromosomal localization of 3 human cytochrome-p450-2c genes (cyp2c8, 2c9 and 2c10) at 10q24.1" *Jpn. J. Hum. Genet.* **39**, 337-343.

14. Cohen, D., Chumakov, I., and Weissenbach, J. (1993) "A first generation physical map of the human genome" *Nature* **366**, 698-701.

15. Parmiter, A.H., Balaban, G., Clark, W.H.J., and Nowell, P.C. (1988) "Possible involvement of the chromosome region 10q24-q26 in early stages of melanocytic neoplasia" *Cancer Genet. Cytogenet.* **30**, 313-317.

16. Ransom, D.T., Ritland, S.R., Moertel, C.A., Dahl, R.J., O'Fallon, J.R., Scheithauer, B.W., Kimmel, D.W., Kelly, P.J., Olopade, O.I., Diaz, M.O., and Jenkins, R.B. (1992) "Correlation of cytogenetic analysis and loss of heterozygosity studies in human diffuse astrocytomas and mixed oligo-atrocytomas" *Genes Chrom.*

Cancer 5, 357-374.

17. Rasheed, B.K.A., Fuller, G.N., Friedman, A.H., Bigner, D.D., and Bigner, S.H. (1992) "Loss of heterozygosity for 10q loci in human gliomas" *Genes Chrom. Cancer* 5, 75-82.

18. Speaks, S.L., Sanger, W.G., Masih, A.S., Harrington, D.S., Hess, M., and Armitage, J.O. (1992) "Recurrent abnormalities of chromosome bands 10q23-q25 in non-Hodgkins lymphoma" *Genes Chrom. Cancer* 5, 239-243.

19. Fults, D., and Pedone, C. (1993) "Deletion mapping of the long arm of chromosome 10 in glioblastoma multiforme" *Genes Chrom. Cancer* 7, 173-177.

20. Karlbom, A.E., James, C.D., Boethius, J., Cavenee, W.K., Collins, V.P., Nordenskjold, M., and Larsson, C. (1993) "Loss of heterozygosity in malignant gliomas involves at least three distinct regions on chromosome 10" *Hum. Genet.* 92, 169-174.

21. Herbst, R.A., Weiss, J., Ehnis, A., Cavanee, W.K., and Arden, K.C. (1994) "Loss of heterozygosity for 10q22-10qter in malignant melanoma progression" *Cancer Res.* 54, 3111-3114.

CLAIMS

1. A nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215 provided that the nucleic acid is not any one of the yeast artificial chromosomes (YACs) 746-H-8, 821-D-2, 831-E-5, 921-F-8, 738-B-12, 796-D-5, 829-E-1, 678-F-1, 839-B-1, 734-B-4, 7B-F12, 757-D-8, 773-C-2, 787-D-7, 831-E-9, 855-D-2, 855-G-4, 876-G-11, 894-H-5, 922-E-6, 934-D-3, 964-A-8, 968-E-6 or 24G-A10 and is not any one of the expressed sequence tags (ESTs) as described in Tables 3 to 22, and is not any one of the bacterial artificial chromosomes (BACs) or P1-derived artificial chromosomes (PACs) B2F20, P40F10, P72G8, P74N2, P274D21, B76I10, B79A19, B7901, B93F12, B122L22, P201J8, P201P5, P209K3, P316N14, B46B12, B60C5, B145C22, B150K4, B150N3, B181F15, and B188L22.

2. A nucleic acid according to Claim 1 capable of selectively hybridising to the region of human chromosome 10 bounded by DNA defined by the markers D10S541 and D10S1765.

3. A nucleic acid according to Claim 1 or Claim 2 capable of selectively hybridising to the human-derived DNA of any one of the yeast artificial chromosomes (YACs) 746-H-8, 821-D-2, 831-E-5, 921-F-8, 796-D-5, 829-E-1, 839-B-1, 734-B-4 or 24G-A10, or of any one of the BACs or PACs B2F20, P40F10, P72G8, P74N2, P274D21, B76I10, B79A19, B7901, B93F12, B122L22, P201J8, P201P5, P209K3, P316N14, B46B12, B60C5, B145C22, B150K4, B150N3, B181F15, and B188L22.

4. A nucleic acid according to Claim 3 wherein the YAC is any one of 921-F-8, 746-H-8, 821-D-2, 831-E-5, 796-D-5, 24G-A-10 and 734-B-4, or the BAC is any one of B2F20, B46B12, B60C5, B150K4, B150N3, B145C22, B181F15 and B188L22, or the PAC is either of P40F10 and P274D21.

5. A nucleic acid according to any one of the preceding claims comprising a gene which corresponds to the cDNA insert of IMAGE clone 264611 or a fragment or variant of said gene.

6. A nucleic acid according to Claim 1 capable of selectively hybridising to the gene as defined in Claim 5.

7. A nucleic acid according to any one of the preceding claims wherein the nucleic acid is DNA.

8. A nucleic acid according to any one of the preceding claims wherein the nucleic acid is single-stranded.

9. A nucleic acid according to any one of the preceding claims wherein the nucleic acid has fewer than 10 000 base pairs when the nucleic acid is double-stranded or 10 000 bases when the nucleic acid is single-stranded.

10. A nucleic acid according to any one of the preceding claims wherein the nucleic acid has fewer than 1000 base pairs when the nucleic acid is double-stranded or 1000 bases when the nucleic acid is single-stranded.

11. A nucleic acid according to any one of the preceding claims

wherein the nucleic acid has from 10 to 100 base pairs when the nucleic acid is double-stranded or from 10 to 100 bases when the nucleic acid is single-stranded.

12. A nucleic acid according to any one of the preceding claims wherein the nucleic acid has from 15 to 30 base pairs when the nucleic acid is double-stranded or from 15 to 30 bases when the nucleic acid is single-stranded.

13. A nucleic acid according to Claim 1 comprising a tumour suppressor gene or fragment or variant thereof.

14. A nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215 comprising a product of a tumour suppressor gene or derivative or fragment or variant thereof.

15. A nucleic acid according to Claim 13 comprising a gene corresponding to the cDNA insert of clone IMAGE 264611 or a fragment or variant of said gene.

16. A nucleic acid according to Claim 13 or 14 capable of selectively hybridising to the gene as defined in Claim 15.

17. A nucleic acid according to Claim 13 wherein the nucleic acid is a cDNA.

18. A nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined

by the markers D10S541 and D10S215, further comprising a detectable label.

19. A nucleic acid according to Claim 18, the nucleic acid comprising the human-derived sequence in any one of the expressed sequence tags (ESTs) as described in Tables 3 to 22 or the intronic sequences in Figures 8 to 15.

20. A nucleic acid according to Claim 18, the nucleic acid comprising a gene corresponding to the cDNA insert of clone IMAGE 264611 or a fragment or variant of said gene.

21. A nucleic acid according to Claim 18 capable of selectively hybridising to the gene as defined in Claim 20.

22. A nucleic acid according to Claim 1 capable of selectively hybridising to the human-derived sequence in any one of the expressed sequence tags (ESTs) as described in Tables 3 to 22, or the intronic sequences of Figures 8 to 15.

23. A nucleic acid according to Claim 22 comprising the gene from which the expressed sequence tag (EST) is derived or the gene which contains said intronic sequence or a fragment or variant thereof.

24. A nucleic acid according to Claim 1 selected from the group consisting of primers suitable for amplifying DNA and which hybridise to the expressed sequence tags (ESTs) or intronic sequences as defined in Claim 19.

25. A method for determining the susceptibility of a patient to cancer comprising the steps of

(i) obtaining a sample containing nucleic acid derived from the patient; and

(ii) contacting the said nucleic acid with a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215.

26. A method of diagnosing cancer in a patient comprising the steps of

(i) obtaining a sample containing nucleic acid derived from the patient; and

(ii) contacting the said nucleic acid with a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215.

27. A method of predicting the relative prospects of a particular outcome of a cancer in a patient comprising the steps of

(i) obtaining a sample containing nucleic acid derived from the patient; and

(ii) contacting the said nucleic acid with a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215.

28. A method according to Claim 25, 26 or 28 wherein the cancer is prostate cancer.

29. A method according to any one of Claims 25 to 28 wherein the said

sample is selected from the group consisting of prostate tissue, blood, semen or urine.

30. A method according to any one of Claims 25 to 29 wherein the nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215 further comprises a detectable label.

31. A method according to any one of Claims 25 to 30 wherein the nucleic acid is capable of selectively hybridising to the region of human chromosome 10 bounded by DNA defined by the markers D10S541 and AFM337xf9.

32. A method according to Claim 31 wherein the nucleic acid is capable of selectively hybridising to the human-derived DNA of any one of the yeast artificial chromosomes (YACs) 746-H-8, 821-D-2, 831-E-5, 921-F-8, 796-D-5, 829-E-1, 839-B-1, 734-B-4 or 24G-A10 or of any one of the BACs or PACs B2F20, P40F10, P72G8, P74N2, P274D21, B76I10, B79A19, B7901, B93F12, B122L22, P201J8, P201P5, P209K3, P316N14, B46B12, B60C5, B145C22, B150K4, B150N3, B181F15, and B188L22.

33. A method according to Claim 32 wherein the YAC is any one of 921-F-8, 746-H-8, 821-D-2, 831-E-5, 796-D-5, 24G-A-10 or 734-B-4 or of any one of the BACs or PACs B2F20, P40F10, P72G8, P74N2, P274D21, B76I10, B79A19, B7901, B93F12, B122L22, P201J8, P201P5, P209K3, P316N14, B46B12, B60C5, B145C22, B150K4, B150N3, B181F15, and B188L22.

34. A method according to Claim 31 wherein the nucleic acid comprises a gene which corresponds to the cDNA insert of IMAGE clone 264611 or a fragment or variant of said gene.

35. A method according to Claim 31 wherein the nucleic acid is capable of selectively hybridising to the gene as defined in Claim 34.

36. A method according to Claim 31 wherein the nucleic acid is DNA.

37. A method according to Claim 31 wherein the nucleic acid is single-stranded.

38. A method according to Claim 31 wherein the nucleic acid has fewer than 10 000 base pairs when the nucleic acid is double-stranded or bases when the nucleic acid is single-stranded.

39. A method according to Claim 31 wherein the nucleic acid has fewer than 1000 base pairs when the nucleic acid is double-stranded or bases when the nucleic acid is single-stranded.

40. A method according to Claim 31 wherein the nucleic acid has from 10 to 100 base pairs when the nucleic acid is double-stranded or bases when the nucleic acid is single-stranded.

41. A method according to Claim 31 wherein the nucleic acid has from 15 to 30 bases pairs when the nucleic acid is double-stranded or bases when the nucleic acid is single-stranded.

42. A method according to Claim 31 wherein the nucleic acid

comprises a tumour suppressor gene or fragment or variant thereof.

43. A method according to Claim 31 wherein the nucleic acid is, or is capable of hybridising to, the human-derived sequence in any one of the expressed sequence tags (ESTs) as described in Tables 3 to 22 or intronic sequences as described in Figures 8 to 15.

44. A method according to Claim 43 wherein the nucleic acid is selected from the group consisting of primers suitable for amplifying DNA from the expressed sequence tags (ESTs) or intronic sequences as defined in Claim 19.

45. A system for detecting the presence or absence of, or mutation in, the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215, the system comprising a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215 and a nucleoside triphosphate or deoxynucleotide triphosphate or derivative thereof.

46. A system according to Claim 45 wherein the nucleoside triphosphate or deoxynucleoside triphosphate is detectably labelled, preferably radioactively labelled.

47. A system for detecting the presence or absence of, or mutation in, the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215, the system comprising a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215 and a nucleic acid

modifying enzyme.

48. A system according to Claim 47 wherein the nucleic acid modifying enzyme is selected from the group consisting of DNA polymerases, DNA ligases, polynucleotide kinases, restriction endonucleases, or nucleases.

49. A polypeptide capable of being encoded by a nucleic acid according to any one of Claims 13 to 17.

50. A molecule capable of binding to a polypeptide according to Claim 49.

51. A method for determining the susceptibility of a patient to cancer comprising the steps of

(i) obtaining a sample containing protein derived from the patient; and

(ii) determining the relative amount in the said sample of the polypeptide according to Claim 49, or whether there is a truncation of, or loss of function of, the polypeptide according to Claim 49.

52. A method of diagnosing cancer in a patient comprising the steps of

(i) obtaining a sample containing protein derived from the patient; and

(ii) determining the relative amount in the said sample of the polypeptide according to Claim 49, or whether there is a truncation of, or loss of function of, the polypeptide according to Claim 49.

53. A method of predicting the relative prospects of a particular outcome of a cancer in a patient comprising the steps of

(i) obtaining a sample containing protein derived from the patient; and

(ii) determining the relative amount in the said sample of the polypeptide according to Claim 49, or whether there is a truncation of, or loss of function of, the polypeptide according to Claim 49.

54. A method according to any one of Claims 51 to 53 wherein the cancer is prostate cancer.

55. A method according to any one of Claims 51 to 53 wherein the polypeptide according to Claim 49 is detected using a molecule according to Claim 50.

56. A method according to any one of Claims 51 to 55 wherein the said sample is selected from the group consisting of prostate tissue, blood, urine or semen.

57. A method according to Claim 55 wherein the molecule comprises a detectable label.

58. The use of a nucleic acid capable of selectively hybridising to the region of chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215 in the manufacture of a reagent for diagnosing cancer; or in the manufacture of a medicament for treating cancer.

59. A method of treating cancer comprising the step of administering

to the patient a nucleic acid capable of selectively hybridising to the region of human chromosome 10 which region is bounded by DNA defined by the markers D10S541 and D10S215.

60. A method according to Claim 59 wherein the nucleic acid is a tumour suppressor gene or a fragment or derivative thereof.

61. A method according to Claim 59 or 60 wherein the nucleic acid comprises a viral vector.

62. A method according to Claim 59 wherein the cancer is prostate cancer.

63. A method of treating cancer comprising the step of administering to the patient a molecule according to Claim 50, the said molecule further comprising a cytotoxic moiety.

64. Use of a molecule according to Claim 50, the said molecule further comprising a cytotoxic moiety, in the manufacture of a medicament for treating cancer.

65. A method of determining loss of heterozygosity in a tissue sample, the method comprising the steps of (i) obtaining a sample containing nucleic acid derived from the tissue and (ii) comparing a microsatellite profile of the said nucleic acid with that of a reference (homozygous) tissue, the microsatellite(s) being chosen by reference to the D10S541-D10S215 interval.

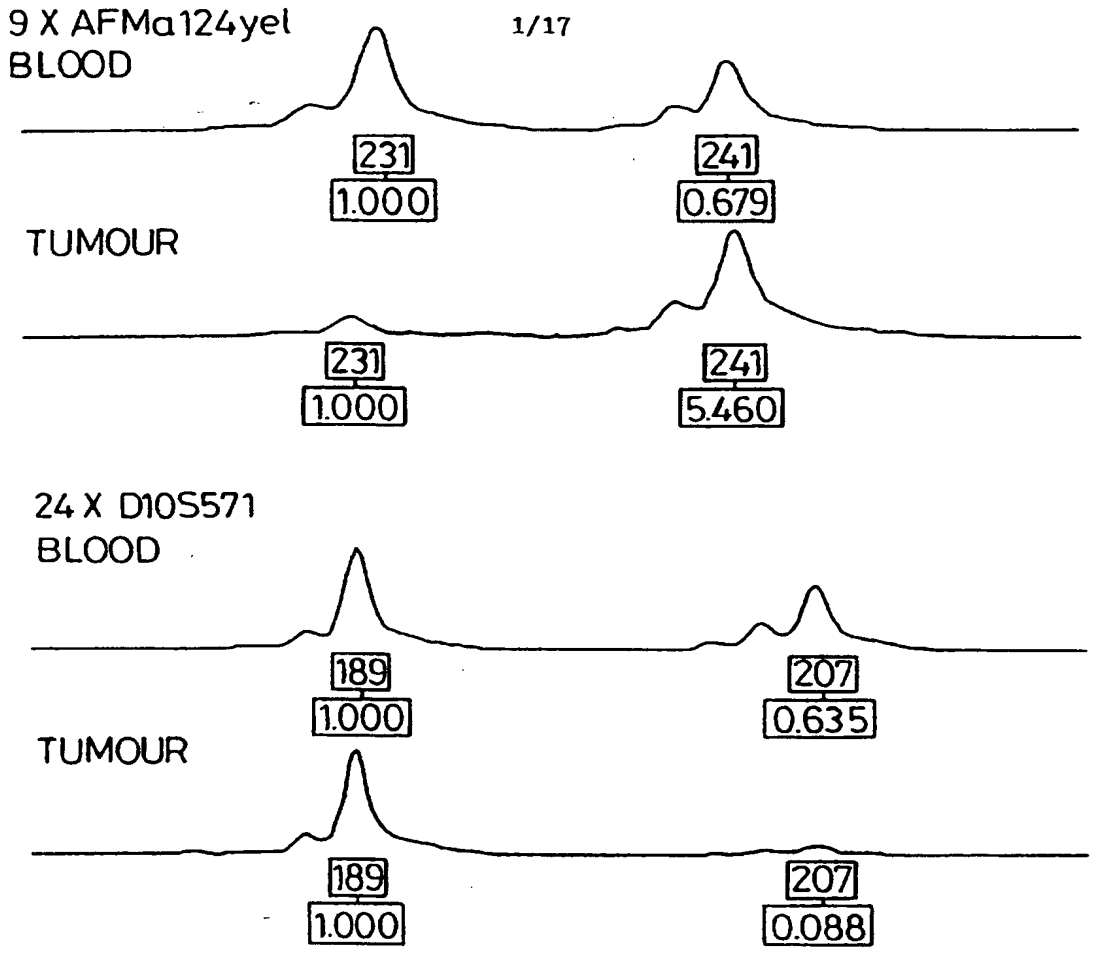

9 X AFMa124yel
BLOOD
231
1.000
241
0.679
TUMOUR
231
1.000
241
5.460
24 X D10S571
BLOOD
189
1.000
207
0.635
TUMOUR
189
1.000
207
0.088

Fig. 1a

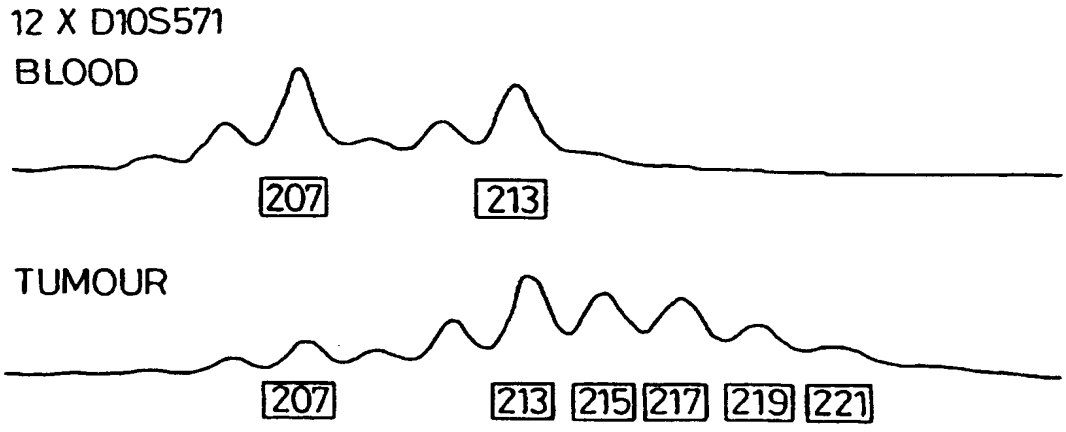

12 X D10S571
BLOOD
207
213
TUMOUR
207
213
215
217
219
221

Fig. 1b

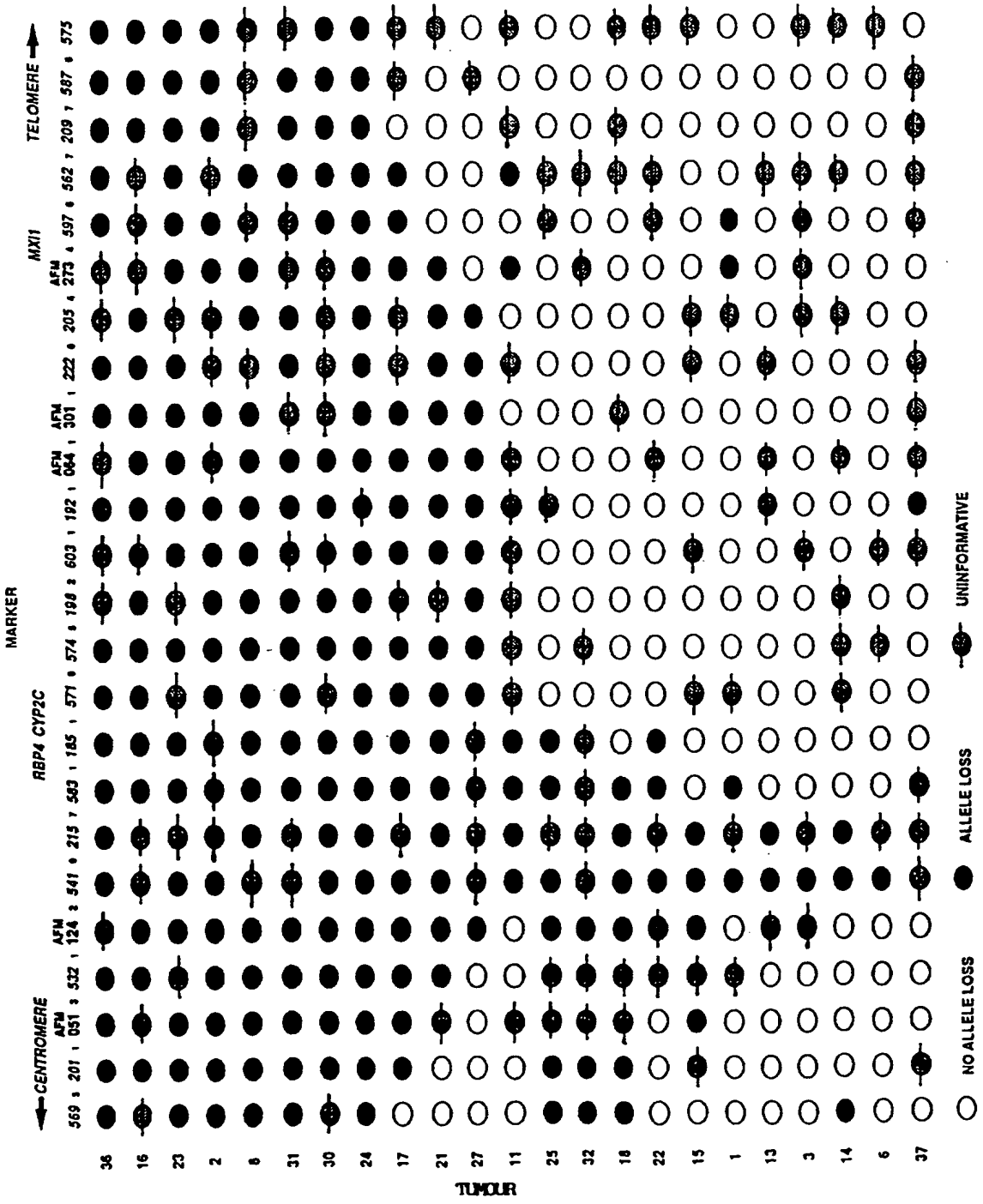

CENTROMERE
TELOMERE
MARKER
RBP4 CYP2C
MXI1
AFM 124 x 532
AFM 273 x 597
AFM 301
AFM 044
569 x 201 x 051 x 532 x 124 x 541 x 215 x 583 x 185 x 571 x 574 x 198 x 603 x 192 x 044 x 301 x 222 x 205 x 273 x 597 x 562 x 209 x 587 x 575
36 16 23 2 8 31 30 24 17 21 27 11 25 32 18 22 15 1 13 3 14 6 37
TUMOUR
UNINFORMATIVE
ALLELE LOSS
NO ALLELE LOSS

FIGURE 2

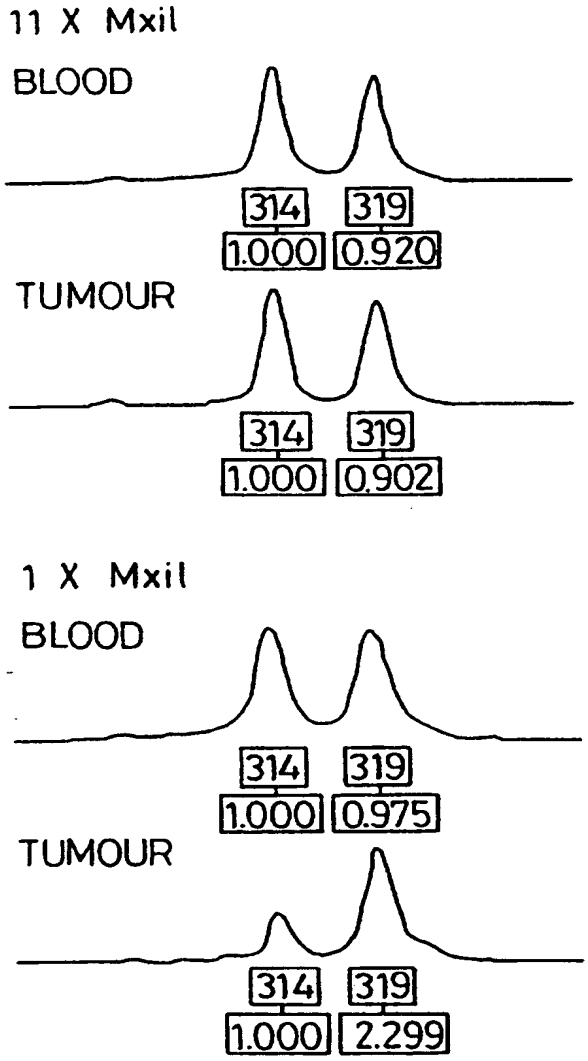

11 X Mxil
BLOOD
314
319
1.000
0.920
TUMOUR
314
319
1.000
0.902
1 X Mxil
BLOOD
314
319
1.000
0.975
TUMOUR
314
319
1.000
2.299

Fig. 3

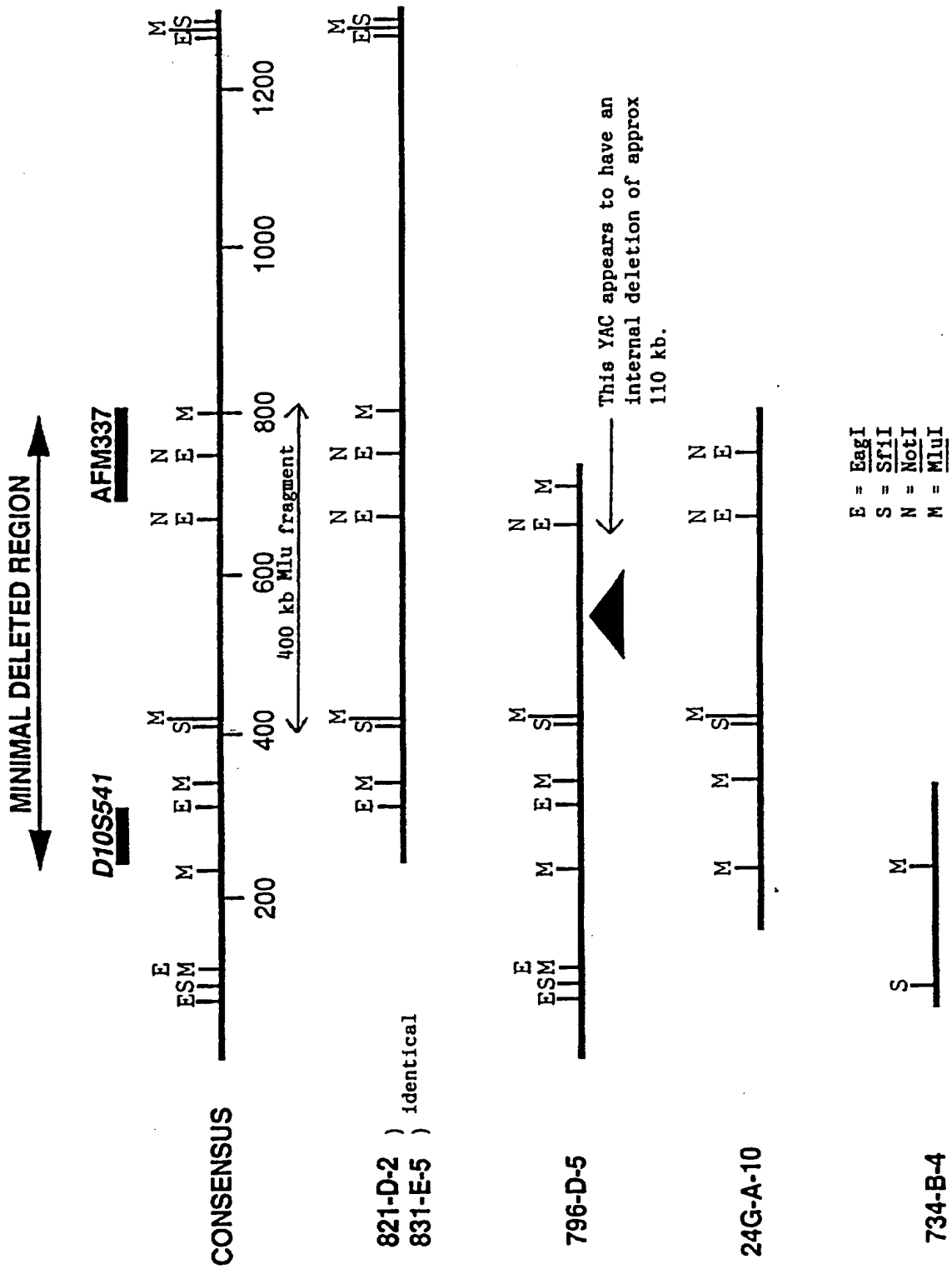

MINIMAL DELETED REGION
D10S541
AFM337
CONSENSUS
821-D-2
831-E-5
identical
796-D-5
24G-A-10
734-B-4
400 kb Mlu fragment
This YAC appears to have an internal deletion of approx 110 kb.
E = EagI
S = SfiI
N = NotI
M = MluI

FIGURE 4(a)

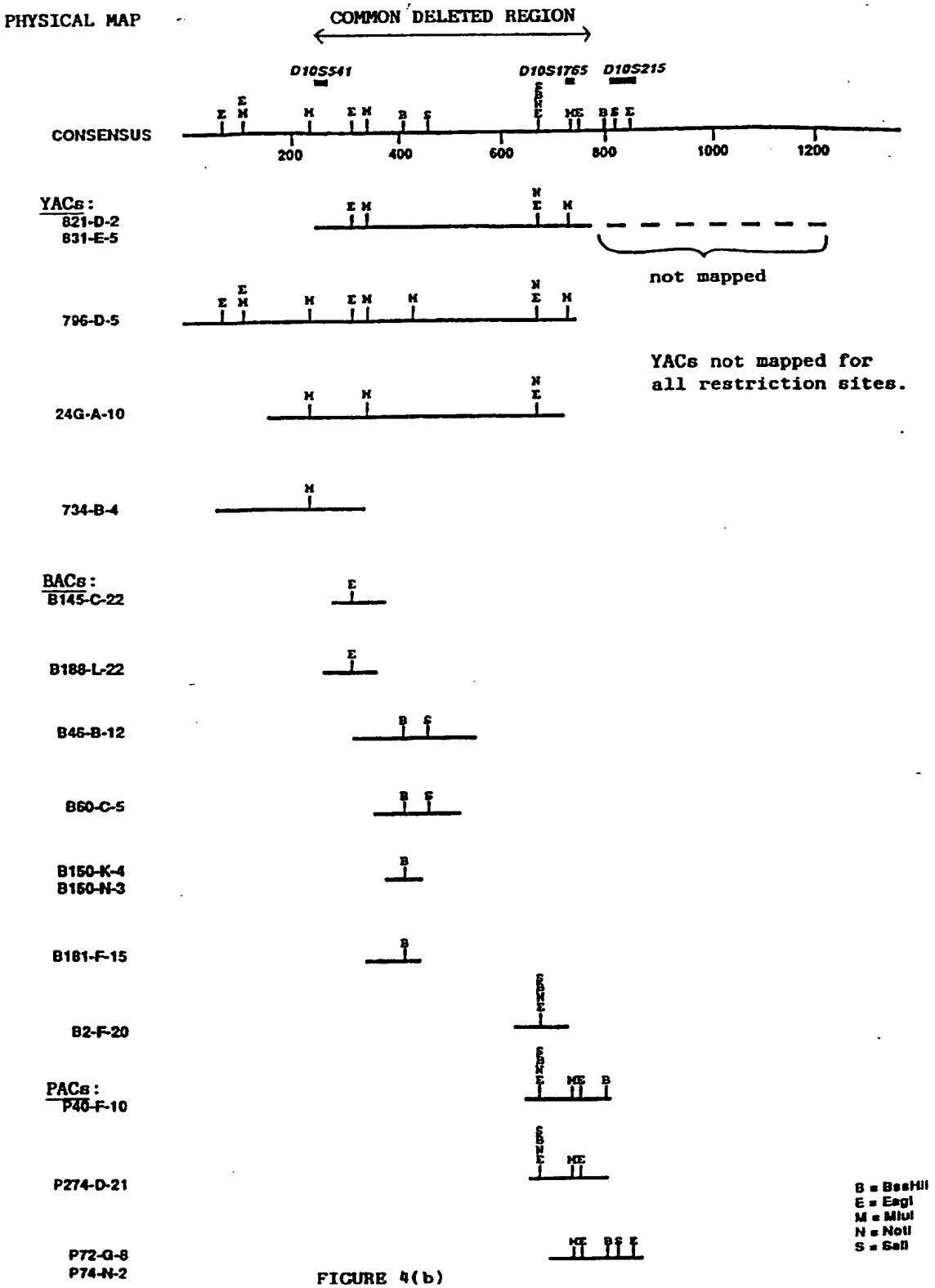

PHYSICAL MAP
COMMON DELETED REGION
D10S541
D10S1765
D10S215
CONSENSUS
200
400
600
800
1000
1200
YACs:
821-D-2
831-E-5
not mapped
796-D-5
YACs not mapped for all restriction sites.
24G-A-10
734-B-4
BACs:
B145-C-22
B188-L-22
B46-B-12
B60-C-5
B150-K-4
B150-N-3
B181-F-15
B2-F-20
PACs:
P46-F-10
P274-D-21
P72-G-8
P74-N-2
B = BssHII
E = EagI
M = MluI
N = NotI
S = SalI

FIGURE 4(b)

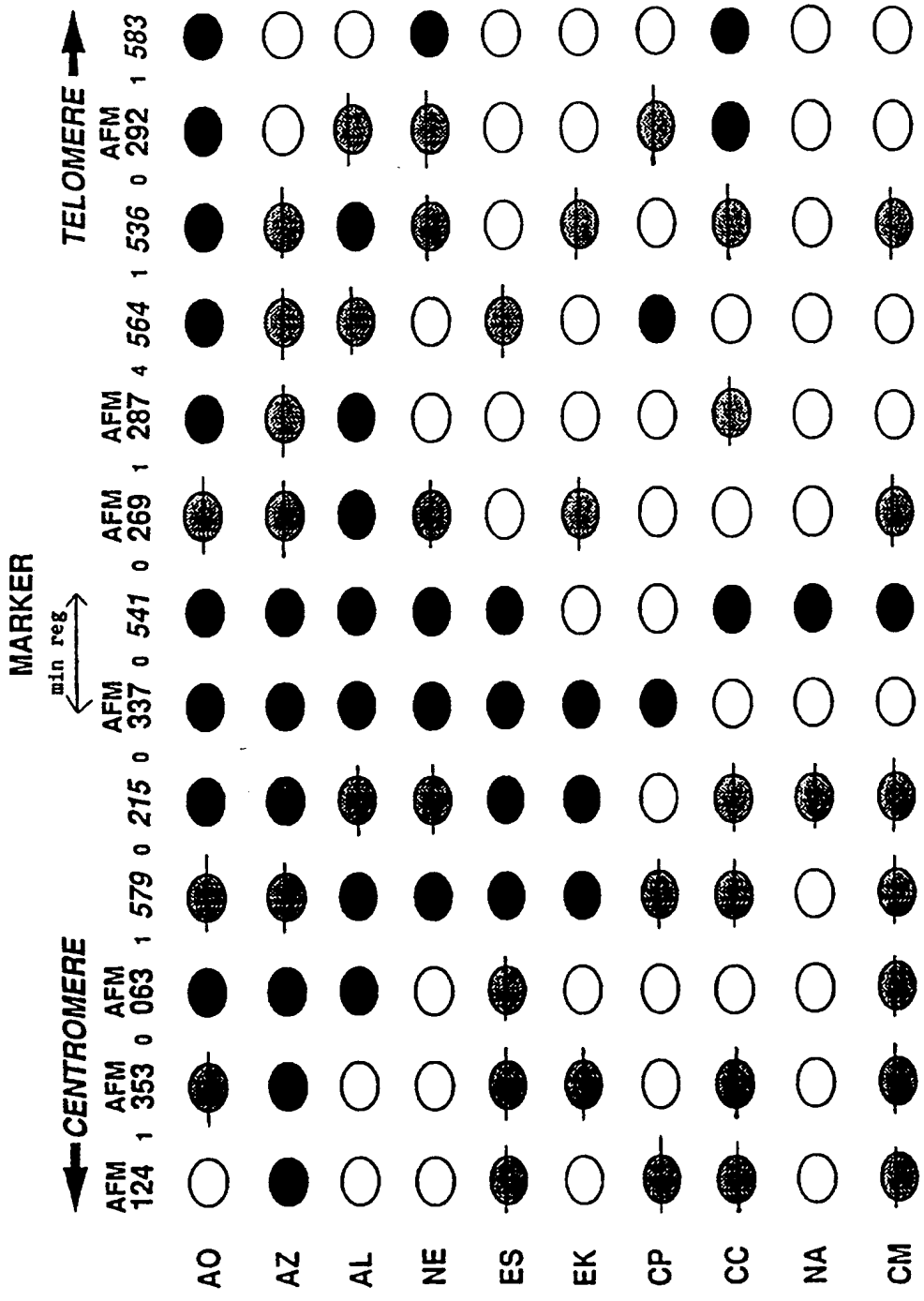

TELOMERE
CENTROMERE
MARKER
min reg
AFM 124 1
AFM 353 0
AFM 063 1
579 0
215 0
AFM 337 0
541 0
AFM 269 1
AFM 287 4
564 1
536 0
AFM 292 1
583
AO
AZ
AL
NE
ES
EK
CP
CC
NA
CM
TUMOUR

FIGURE 5

```
   1  CGGCCGCGGC GGCTGCAGCT CCAGGGAGGG GGTCTGAGTC GCCTGTCACC
  51  ATTTCCAGGG CTGGGAACGC CGGAGAGTTG GTCTCTCCCC TTCTACTGCC
 101  TCCAACACGG CGGCGGCGGC GGCGGCACAT CCAGGGACCC GGGCCGGTTT
 151  TAAACCTCCC GTCCGCCGCC GCCGCACCCC CCGTGGCCCG GGCTCCGGAG
 201  GCCGCCGGCG GAAGCAGCCG TTCGGAGGAT TATTCGTCTT CTCCCCATTC
 251  CGCTGCCGCC GCTGCCAGGC CTCTGGCTGC TGAGGAGAAG CAGGCCCAGT
 301  CGCTGCAACC ATCCAGCAGC CGCCGCAGCA GCCATTACCC GGCTGCGGTC
 351  CAGAGCCAAG CGGCGGCAGA GCGAGGGGCA TCAGCTACCG CCAAGTCCAG
 401  AGCCATTTCC ATCCTGCAGA AGAAGCCCCG CCACCAGCAG CTTCTGCCAT
 451  CTCTCTCCTC CTTTTTCTTC AGCCACAGGC TCCCAGACAT GACAGCCATC
 501  ATCAAAGAGA TCGTTAGCAG AAACAAAAGG AGATATCAAG AGGATGGATT
                      ss
 551  CGACTTAGAC TTGACCTATA TTTATCCAAA CATTATTGCT ATGGGATTTC
 601  CTGCAGAAAG ACTTGAAGGC GTATACAGGA ACAATATTGA TGATGTAGTA
                                                    ss
 651  AGGTTTTTGG ATTCAAAGCA TAAAAACCAT TACAAGATAT ACAATCTTTG
                                          ss
 701  TGCTGAAAGA CATTATGACA CCGCCAAATT TAATTGCAGA GTTGCACAAT
 751  ATCCTTTTGA AGACCATAAC CCACCACAGC TAGAACTTAT CAAACCCTTT
 801  TGTGAAGATC TTGACCAATG GCTAAGTGAA GATGACAATC ATGTTGCAGC
 851  AATTCACTGT AAAGCTGGAA AGGGACGAAC TGGTGTAATG ATATGTGCAT
 901  ATTTATTACA TCGGGGCAAA TTTTTAAAGG CACAAGAGGC CCTAGATTTC
                           ss
 951  TATGGGGAAG TAAGGACCAG AGACAAAAAG GGAGTAACTA TTCCCAGTCA
1001  GAGGCGCTAT GTGTATTATT ATAGCTACCT GTTAAAGAAT CATCTGGATT
1051  ATAGACCAGT GGCACTGTTG TTTCACAAGA TGATGTTTGA AACTATTCCA
                      ss
1101  ATGTTCAGTG GCGGAACTTG CAATCCTCAG TTTGTGGTCT GCCAGCTAAA
1151  GGTGAAGATA TATTCCTCCA ATTCAGGACC CACACGACGG GAAGACAAGT
1201  TCATGTACTT TGAGTTCCCT CAGCCGTTAC CTGTGTGTGG TGATATCAAA
                                          ss
1251  GTAGAGTTCT TCCACAAACA GAACAAGATG CTAAAAAAGG ACAAAATGTT
1301  TCACTTTTGG GTAAATACAT TCTTCATACC AGGACCAGAG GAAACCTCAG
1351  AAAAAGTAGA AAATGGAAGT CTATGTGATC AAGAAATCGA TAGCATTTGC
1401  AGTATAGAGC GTGCAGATAA TGACAAGGAA TATCTAGTAC TTACTTTAAC
1451  ARAAAATGAT CTTGACAAAG CAAATAAAGA CAAAGCCAAC CGATACTTTT
                ss
1501  CTCCAAATTT TAAGGTGAAG CTGTACTTCA CAAAAACAGT AGAGGAGCCG
```

FIGURE 6 (1 of 2)

```
1551  TCAAATCCAG AGGCTAGCAG TTCAACTTCT GTAACACCAG ATGTTAGTGA
1601  CAATGAACCT GATCATTATA GATATTCTGA CACCACTGAC TCTGATCCAG
1651  AGAATGAACC TTTTGATGAA GATCAGCATA CACAAATTAC AAAAGTCtga
1701  attttttttt atcaagaggg ataaaacacc atgaaaataa acttgaataa
1751  actgaaaaaa aaaaaaaaaa aaa
```

FIGURE 6 (2 of 2)

TRANSLATION

```
  1  AAAAAAPGRG SESPVTISRA GNAGELVSPL LLPPTRRRRR RHIQGPGPVL
 51  NLPSAAAAPP VARAPEAAGG SSRSEDYSSS PHSAAAAARP LAAEEKQAQS
101  LQPSSSRRSS HYPAAVQSQA AAERGASATA KSRAISILQK KPRHQQLLPS
151  LSSFFFSHRL PDMTAIIKEI VSRNKRRYQE DGFDLDLTYI YPNIIAMGFP
201  AERLEGVYRN NIDDVVRFLD SKHKNHYKIY NLCAERHYDT AKFNCRVAQY
251  PFEDHNPPQL ELIKPFCEDL DQWLSEDDNH VAAIHCKAGK GRTGVMICAY
301  LLHRGKFLKA QEALDFYGEV RTRDKKGVTI PSQRRYVYYY SYLLKNHLDY
351  RPVALLFHKM MFETIPMFSG GTCNPQFVVC QLKVKIYSSN SGPTRREDKF
401  MYFEFPQPLP VCGDIKVEFF HKQNKMLKKD KMFHFWVNTF FIPGPEETSE
451  KVENGSLCDQ EIDSICSIER ADNDKEYLVL TLTXNDLDKA NKDKANRYFS
501  PNFKVKLYFT KTVEEPSNPE ASSSTSVTPD VSDNEPDHYR YSDTTDSDPE
551  NEPFDEDQHT QITKV*IFFY QEG*NTMKIN LNKLKKKKKK
```

FIGURE 7

EXON 1

```
   1  CGGCCGCGGC GGCTGCAGCT CCAGGGAGGG GGTCTGAGTC GCCTGTCACC
  51  ATTTCCAGGG CTGGGAACGC CGGAGAGTTG GTCTCTCCCC TTCTACTGCC
 101  TCCAACACGG CGGCGGCGGC GGCGGCACAT CCAGGGACCC GGGCCGGTTT
 151  TAAACCTCCC GTCCGCCGCC GCCGCACCCC CCGTGGCCCG GGCTCCGGAG
 201  GCCGCCGGCG GAAGCAGCCG TTCGGAGGAT TATTCGTCTT CTCCCCATTC
 251  CGCTGCCGCC GCTGCCAGGC CTCTGGCTGC TGAGGAGAAG CAGGCCCAGT
 301  CGCTGCAACC ATCCAGCAGC CGCCGCAGCA GCCATTACCC GGCTGCGGTC
 351  CAGAGCCAAG CGGCGGCAGA GCGAGGGGCA TCAGCTACCG CCAAGTCCAG
 401  AGCCATTTCC ATCCTGCAGA AGAAGCCCCG CCACCAGCAG CTTCTGCCAT
 451  CTCTCTCCTC CTTTTTCTTC AGCCACAGGC TCCCAGACAT GACAGCCATC
 501  ATCAAAGAGA TCGTTAGCAG AAACAAAAGG AGATATCAAG AGGATGGATT
 551  CGACTTAGAC TTGACCTgta tccatttctg cggctgctcc tctttacctt
 601  tctgtcactc tcttagaacg tgggagtaga cggatgcgaa aatgtccgta
 651  gtttgggtga ctataacatt taaccctggt caggttgcta ggtcatatat
 701  tttgtgtttc ctttctgtgt attcaaccta gggtgtgttt ggctagacgg
 751  aactcttgcc tggttgcaag tgtcaagcca ccgattgctt tcttaggcta
 801  tctatatggt ctcttcctga gggctattgt ccgttaatac agaatacagt
 851  aaggagagga cagcgatcct a
```

FIGURE 8

EXON 2

```
   1  tcgnatccnt acccgttcgt acgagaatcg ctgtccctct cccttctaat
  51  gttttaaaaa gtattctttt agtttgattg ctgcatattt cagatatttn
 101  ctttccttaa ctaaagtaac tcagATATTT ATCCAAACAT TATTGCTATG
 151  GGATTTCCTG CAGAAAGACT TGAAGGCGTA TACAGGAACA ATATTGATGA
 201  TGTAGTAAGG TTTTTGGATT CAAAGCATAA AAACCATTAC AAGATATACA
 251  ATCT
```

FIGURE 9

EXON 3

```
   1  taaaacacag cataatatgt gtcacattat aaagattcag gcaatgtttg
  51  ttagtattag tacttttttt tcttcctaag tgcaaaagat aactttatat
 101  cacttttaaa cttttctttt agTTGTGCTG AAAGACATTA TGACACCGCC
 151  AAATTTAATT GCAGaggtag gtatgaatgt actgtactat gttgtataac
 201  ttaaacccga tagactgtat cttactgtca taacaataat gagtcatcca
 251  gattatcgag tgagatacat atttatctta agaattatct ttaaaaattt
 301  caaaaatttt aatttgactg ttgtgtttta ggaaaaagta ttgcataaag
 351  ctattaatat tgtcaggaag actaaagtgc agcata
```

FIGURE 10

EXON 4

```
   1  ttttctacct ctaatngctg acntatgcta ccagtccgta tagcgtaaat
  51  tcccagaata tatcctcctg aataaaatgg gggaaaataa tacctggctt
 101  ccttaatgat tatatttaan acttatcaan anactatttt ctatttaaca
 151  attagaaagt taagcaatac attatttttc tctggaatcc agtgtttctt
 201  ttaaatacct gttaagtttg tatgcaacat ttctaaagtt acctacttgt
 251  taattaaaaa ttcaanagtt ttttttnctt attctgaggt tatcttttta
 301  ccacAGTTGC ACAATATCCT TTTGAAGACC ATAACCCACC ACAGCTAGAA
 351  CTTATCAAAC CCTTTTGTGA AGATCTTGAC CAATGGCTAA GTGAAGATGA
 401  CAATCATGTT GCAGCAATTC ACTGTAAAGC TGGAAAGGGA CGAACTGGTG
 451  TAATGATATG TGCATATTTA TTACATCGGG GCAAATTTTT AAAGGCACAA
 501  GAGGCCCTAG ATTTCTATGG GGAAGTAAGG ACCAGAGACA AAAaaggtaa
 551  gttatttttt gatgtttttc ctttcctctt cctggatctg agaatttatt
 601  ggaaaacaga ttttgggttt ctttttttct tcagttttat tgaggtgtaa
 651  ttgcacaagt aaaaattata tataaataca atgtataata tgatgtttgg
 701  atgtatgtgt atatacattg tgaa
```

FIGURE 11

EXON 5

```
   1  aaggtcaaat gtctaatgta tatatgttct taaatggcta cgacccagtt
  51  accatagcaa tttagtgaaa taactataat ggaacatttt ttttcaattt
 101  ggcttctctt ttttttctgt ccaccAGGGA GTAACTATTC CCAGTCAGAG
 151  GCGCTATGTG TATTATTATA GCTACCTGTT AAAGAATCAT CTGGATTATA
 201  GACCAGTGGC ACTGTTGTTT CACAAGATGA TGTTTGAAAC TATTCCAATG
 251  TTCAGTGGCG GAACTTGCAg taagtgcttg aatctcatcc ttccatgtta
 301  ttgggaacag ttttcttaac catatctaga agtttacata aaaatttaga
 351  aagaaattta ccacatttga aatttatgca ggagactata tttctgaagc
 401  atttgaacaa attaattagc tttgttgttc aactcattgg gctaaagaag
 451  ccaaaagcaa tgggttttaa tgtagtcgaa gccaaattat atttatgaaa
 501  gaaatattct gtgttataac caccaaatac agcccaattc tg
```

FIGURE 12

EXON 6

```
   1  actctgccac tagaagtcta attttgggac ttactattca tgaaatagga
  51  attgactttn atataagtaa taggacctta ttttgaaggt tcaaactgga
 101  gaaaatctta cattgtttat atttttattt catttanttc agttgatttg
 151  cttgagatca agattgcaga tacagaatcc atatttcgtg tatattgctg
 201  atattaatca ttaaaatcgt ttttgacagt ttgacagtta aaggcatttc
 251  cctgtgaaat aatactggta tgtatttaac catgcagATC CTCAGTTTGT
 301  GGTCTGCCAG CTAAAGGTGA AGATATATTC CTCCAATTCA GGACCCACAC
 351  GACGGGAAGA CAAGTTCATG TACTTTGAGT TCCCTCAGCC GTTACCTGTG
 401  TGTGGTGATA TCAAAGTAGA GTTCTTCCAC AAACAGAACA AGATGCTAAA
 451  AAgggtttgt actttacttt cattgggaga aatatccaaa ataaggacag
 501  attanaagct ntattntatt ttatgacatg taaggaacta taatttgttt
 551  tctattagat ctgccaggtg ttttgcttac tctggcattg gtgagacatt
 601  atangggtaa ataatcctgt ttgaaggaan aggcctat
```

FIGURE 13

EXON 7

```
   1  tttatcttag atcttgtgag attgtatttt tggtttaaaa tttgagaatt
  51  tgagtgaaga aaaatcatgt gaatgaaaat gcaacagata actcagattg
 101  ccttataata gtctttgtgt ttacctttat tcagaatatc aaatgatagt
 151  ttattttgtt gacttttttgc aaatgtttaa cataggtgac agatttnctt
 201  ttttaaaaaa ataaaacatc attaattaaa tatgtcattt catttctttt
 251  tcttttcttt ttttttttt tAGGACAAAA TGTTTCACTT TTGGGTAAAT
 301  ACATTCTTCA TACCAGGACC AGAGGAAACC TCAGAAAAAG TAGAAAATGG
 351  AAGTCTATGT GATCAAGAAA TCGATAGCAT TTGCAGTATA GAGCGTGCAG
 401  ATAATGACAA GGAATATCTA GTACTTACTT TAACARAAAA TGATCTTGAC
 451  AAAGCAAATA AAGACAAAGC CAACCGATAC TTTTCTCCAA ATTTTAAGgt
 501  cagttaaatt aaacattttg tgggggntgg tgacttgtat gtatgtgatg
 551  tgtgtttaat tctaggagta cagaaggaga ggacagcgat
```

FIGURE 14

EXON 8

```
   1  ggaggcagag gttgcagtga gccaagatca tgccactgca ctccagcttg
  51  gcaacagagc aagactcttg tctccagaaa taaaaataaa taaattgtat
 101  taacatcctg atagtttatc tgtttagtac ctagcaagaa agaaaatgtt
 151  gaacatctta agaagagggt catttaaaag gcctcttaaa gatcatgttt
 201  gttacagtgc ttaaaaatta atatgttcat ctgcaaaatg gaataaaaaa
 251  tctgttaaaa atatatttca ctaaatagtt aagatgagtc atatttgtgg
 301  gttttcattt taaatttttct ttctctaGTG AAGCTGTACT TCACAAAAAC
 351  AGTAGAGGAG CCGTCAAATC CAGAGGCTAG CAGTTCAACT TCTGTAACAC
 401  CAGATGTTAG TGACAATGAA CCTGATCATT ATAGATATTC TGACACCACT
 451  GACTCTGATC CAGAGAATGA ACCTTTTGAT GAAGATCAGC ATACACAAAT
 501  TACAAAAGTC tgaatttttt tttatcaaga gggataaaac accatgaaaa
 551  taaacttgaa taaactgaaa aaaaaaaaaa aaaaaa
```

FIGURE 15

INTERNATIONAL SEARCH REPORT

Inte onal Application No
PCT/GB 96/02588

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12Q1/68 C07K14/47

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 95 22624 A (UNIV CALIFORNIA) 24 August 1995 see claims see the whole document ---	1,2, 5-31, 34-65
Y	NATURE GENETICS, vol. 7, June 1994, pages 246-339, XP000644432 GYAPAY ET AL: "The 1993-94 genethon human genetic linkage map" cited in the application see pp 246-249 and 292-295 --- -/--	1,2, 5-31, 34-65

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 February 1997	06.03.97

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+ 31-70) 340-3016	Hagenmaier, S

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 96/02588

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	ONCOGENE, vol. 9, 1994, pages 2999-3003, XP000644763 GAO ET AL.: "High frequency of mutator phenotype in human prostatic adenocarcinoma" cited in the application see the whole document	1,2, 5-31, 34-65
Y	GENES, CHROMOSOMES AND CANCER, vol. 14, October 1995, pages 112-119, XP000645023 NIHEI ET AL.: "Localization of metastasis suppressor gene(s) for rat prostatic cancer to the long arm of human chromosome 10" see the whole document	1,2, 5-31, 34-65
Y	DATABASE EMBL EST, Homo sapiens cDNA clone 126556., 3 April 1995 HILLIER ET AL.: "WashU-Merck EST project" XP002025876 & Accession R06763	1,2, 5-31, 34-65
Y	DATABASE EMBL EST, Homo sapiens cDNA clone 126556 5'., 3 April 1995 HILLIER ET AL.: "WashU-Merck EST project" XP002025965 & Accession R06814	1,2, 5-31, 34-65
Y	DATABASE EMBL EST03045 Homo sapiens cDNA clone HFBCS42, 30 June 1993 ADAMS ET AL: "3400 expressed sequence tags identify diversity of transcripts from human brain" XP002025966 & Accession T05157	1,2, 5-31, 34-65
Y	DATABASE EMBL EST, Homo sapiens cDNA clone 81420 5'., 9 February 1995 HILLIER ET AL.: "WashU-Merck EST project" XP002025967 & Accession T60214	1,2, 5-31, 34-65

-/--

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte onal Application No
PCI/GB 96/02588

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9522624 A	24-08-95	EP 0757723 A	12-02-97
CA 2013430 A	29-09-91	NONE	
WO 9424308 A	27-10-94	AU 6678894 A	08-11-94
WO 9532214 A	30-11-95	AU 2603495 A	18-12-95
		CA 2170627 A	30-11-95
		EP 0714401 A	05-06-96
		FI 961723 A	19-04-96
		NO 961539 A	18-04-96

INTERNATIONAL SEARCH REPORT

Inte onal Application No
PCT/GB 96/02588

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CA 2 013 430 A (LITT (US) AND LUTY (US)) 29 September 1991 see the whole document	1,2, 5-31, 34-65
A	WO 94 24308 A (RAGGIO ITALGENE S P A ;RUSSO GIANDOMENICO (IT); VIRGILIO LAURA (US) 27 October 1994 see the whole document	1,2, 5-31, 34-65
A	NATURE, vol. 359, 29 October 1992, pages 794-801, XP002010900 WEISSENBACH ET AL.: "A second generation linkage map of the human genom" see the whole document	1,2, 5-31, 34-65
A	NATURE, vol. 366, 16 December 1993, pages 698-701, XP002025873 COHEN ET AL.: "A first generation physical map of the human genome" cited in the application see the whole document	1,2, 5-31, 34-65
P,X	CANCER RESEARCH, vol. 55, 1 November 1995, pages 4800-4803, XP002025874 GRAY ET AL.: "Loss of the chromosomal region 10q23-25 in prostate cancer" see the whole document	1,2, 5-31, 34-65
P,Y	THE AMERICAN JOURNAL OF HUMAN GENETICS, vol. 57, no. 4, 24 October 1995, pages A65-346, XP000644590 GRAY ET AL.: "Narrowing a region of prostate tumour suppression at the chromosomal location 10q23-q25" see the whole document	1,2, 5-31, 34-65
P,Y	WO 95 32214 A (CANJI INC ;BOOKSTEIN ROBERT (US); ISAACS WILLIAM B (US)) 30 November 1995 see the whole document	1,2, 5-31, 34-65

3